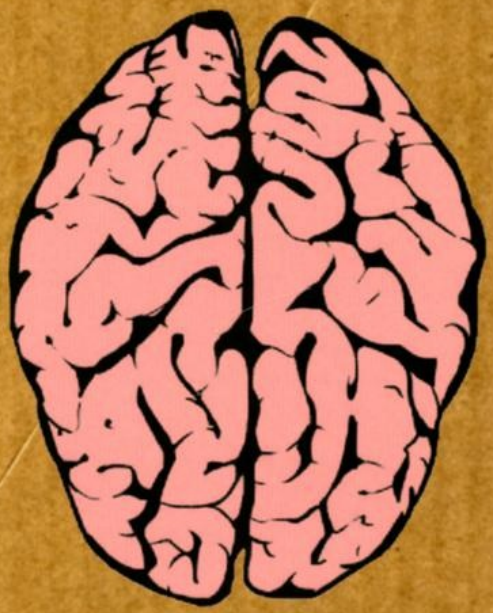
Philip Johnson-Laird

Cómo razonamos

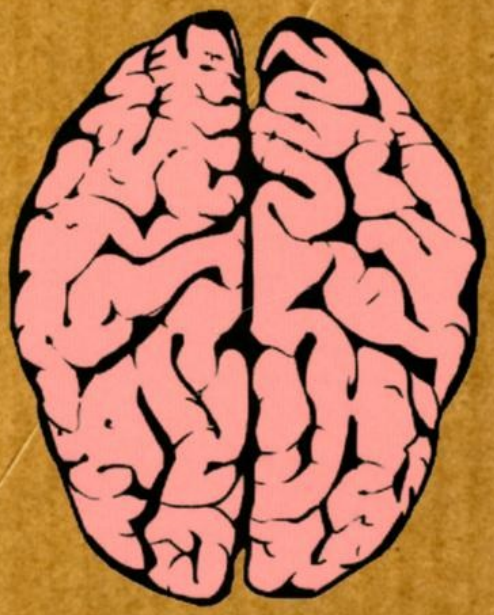
Antonio Machado Libros



Philip Johnson-Laird

Cómo razonamos

Antonio Machado Libros



Philip Johnson-Laird

Machado Nuevo Aprendizaje

Volumen 10 de la colección Machado Nuevo Aprendizaje

Título original: How We Reason

© Oxford University Press, 2006

© De la traducción: M.ª José Pozo Municio

© De la presente edición:

MACHADO GRUPO DE DISTRIBUCIÓN, S.L.

C/ Labradores, 5. Parque Empresarial Prado del Espino

28660 Boadilla del Monte (Madrid)

machadolibros@machadolibros.com

www.machadolibros.com

ISBN: 978-84-9114-243-0

Índice

Prólogo

1. Introducción

Parte I. El mundo en la mente consciente

2. Iconos e imágenes

3. Modelos de posibilidades: de los trucos de magia a las catástrofes

Parte II. El mundo en la mente inconsciente

4. La arquitectura mental y el inconsciente

5. Intuiciones y razonamiento inconsciente

6. Las emociones como inferencias

7. El razonamiento en las enfermedades psicológicas

Parte III. Cómo hacemos deducciones

8. Únicamente conexiones

9. Soy mi propio abuelo: el razonamiento sobre identidades y otras relaciones

10. Los silogismos y el razonamiento sobre propiedades

11. ¿No somos todos optimistas? El caso del razonamiento complejo

Parte IV. Cómo hacemos inducciones

12. La modulación: un paso hacia la inducción

13. El conocimiento y las inducciones

14. El método de Sherlock Holmes: la abducción

15. El balance de probabilidades

Parte V. Qué nos hace racionales

16. Los contraejemplos

17. Verdades, mentiras y razonamiento superior

Parte VI. Cómo desarrollamos la capacidad de razonar

18. Sobre el desarrollo

19. Estrategias y culturas

20. Cómo podemos mejorar nuestro razonamiento

Parte VII. Conocimiento, creencias y problemas

21. Los misterios del si

22. Causas y obligaciones

23. Creencias, herejías y cambios de opinión

24. Cómo resolvemos los problemas

Parte VIII. Razonamiento experto en tecnología, lógica y ciencia

25. Bicicletas voladoras: cómo inventaron el aeroplano los hermanos Wright

26. Desvelar un enigma

27. Sobre el modo de transmisión del cólera

28. Cómo razonamos

Glosario

Notas en los capítulos

Agradecimientos

Referencias

Prólogo

¿Por qué hay quien es un experto razonador y quien es un zoquete? ¿Qué es lo que provoca errores en nuestro razonamiento que pueden llevar a catástrofes como la de Chernobyl? ¿Cómo podemos mejorar nuestro razonamiento? ¿Cómo es posible que tengamos opiniones inconsistentes? ¿Cómo pueden los prestidigitadores inventar ilusiones que nos confunden? ¿Por qué modificamos algunas creencias pero nos aferramos a otras casi sin esfuerzo? ¿Por qué hay quien desarrolla temores irracionales? ¿Por qué los terroristas desarrollan ideologías "peligrosas"? ¿Por qué cree la gente que existen las brujas, que la Tierra es plana y que solo tiene cuatro mil años de existencia (y eso aquí en Estados Unidos)? ¿Por qué en los sistemas legales reales se producen paradojas similares a Trampa-22? ¿No somos racionales? Y, si lo somos, ¿cómo es posible si nunca nos han enseñado lógica?

Este libro aborda todas estas cuestiones, porque todas dependen de cómo razonamos –de cómo nuestras mentes utilizan la información para concluir algo nuevo–. De esto es de lo que trata este libro. Nuestra destreza es la base de nuestra inteligencia. Nuestros errores son la base de nuestra estupidez. El razonamiento correcto puede llevar al éxito y a una vida más larga, y el razonamiento incorrecto puede llevar a la catástrofe. Sin embargo, debido a que gran parte de la vida mental se produce fuera de nuestra consciencia, el saber cómo razonamos y por qué cometemos errores no es algo obvio. A pesar de estas bases inconscientes, en las últimas décadas se ha desarrollado un conocimiento científico sobre el razonamiento, pero ese conocimiento tiene todavía que llegar a quienes no pertenecen al estrecho margen de investigadores de la ciencia cognitiva. Este libro intenta que esa nueva ciencia resulte accesible para todos.

Razonamos de muchas formas, y la mayoría de ellas se describen en este libro.

Razonamos cuando tenemos una intuición. Examinamos las fotos de dos candidatos en unas elecciones, y podemos ver que uno parece más competente que el otro. El proceso va más allá de la simple percepción. Es inferencial, pero no podemos decir qué es lo que da lugar a esa opinión. Incluso podemos razonar sin llegar a ser conscientes de nuestra conclusión –un proceso inconsciente que

los psicólogos detectan solamente en nuestra conducta posterior.

Razonamos cuando tenemos una reacción emocional. Vemos colapsarse las Torres Gemelas y brota en nosotros una emoción. La transición a la emoción es inconsciente, y a veces su fuerza nos sorprende. Un taxista de Nueva York fue testigo de los sucesos del 11 de septiembre de 2001 y cogió tanto terror a los túneles que ya no pudo seguir conduciendo por ellos a los pasajeros. Algunos especialistas creen que la causa de este tipo de reacción neurótica es un razonamiento defectuoso. Esa idea condujo a una revolución en psicoterapia. Pero, cuando mis colegas y yo valoramos el razonamiento de sujetos con tendencias neuróticas, el resultado fue sorprendente. Razonaban mejor que el resto de nosotros, aunque solo sobre temas relacionados con sus síntomas. Este libro describe cómo este dato condujo a una nueva teoría sobre las enfermedades psicológicas y el papel que en ellas desempeña el razonamiento.

Razonamos cuando deducimos las consecuencias lógicas de las premisas. Los sudokus, que se han extendido desde Japón a todo el mundo, se resuelven mediante deducciones lógicas. De Este a Oeste, todos tenemos la capacidad para hacerlas. Quienes son buenos razonadores obtienen mejores resultados en las escuelas y universidades que quienes son menos buenos. Esta es una buena noticia para el mundo académico –debe ser que el estudio y la investigación exigen una capacidad de razonar–. Y es una buena noticia para los test mentales –debe ser que miden algo útil.

Pero es una mala noticia para el puñado de estudiosos que creen que la deducción está fuera de nuestro alcance –que utilizamos otros métodos basados en conjeturas, y que a lo sumo es una destreza que adquirimos solo si estudiamos lógica–. ¡Dígale eso a alguien que hace sudokus! Todos podemos hacer deducciones desde la infancia. Pero, no nos apoyamos en las leyes de la lógica. En lugar de ello, según este libro, razonamos pensando en lo que es posible, y viendo qué tienen en común las posibilidades. Esta teoría predice que deberíamos equivocarnos con ciertos tipos de deducción, y lo hacemos. Estos errores son tan seductores que casi todo el mundo los comete. Hay ilusiones cognitivas en las que no pensamos en las posibilidades correctas lo mismo que cuando los prestidigitadores en el escenario nos enredan con sus trucos.

Razonamos cuando vamos más allá de la información que se nos da. Hacemos una inducción –una generalización, digamoslo así, que nos conduce desde uno o dos hechos a una conclusión sobre todos los hechos del mismo tipo–. Una

experiencia con un cepo es suficiente para que infiramos que en Londres no debemos aparcar donde haya dobles líneas amarillas. Utilizamos nuestro conocimiento para extraer esta conclusión. También usamos nuestro conocimiento cuando inferimos una explicación para un hecho. Como verá, el doctor que sirvió de modelo real para crear a Sherlock Holmes era un maestro en este tipo de razonamiento. Pero, todos nosotros somos capaces de hacerlo.

Razonamos sobre nuestro propio razonamiento. Esta habilidad es necesaria si tenemos que crear la lógica como una disciplina intelectual. Nos ayuda a desarrollar estrategias para razonar sobre problemas que son nuevos para nosotros. La utilizamos cuando tratamos de resolver nuestro primer problema de ingeniería inversa –para entender, por ejemplo, cómo conectar un circuito eléctrico que controle las luces de nuestro recibidor–. Y este libro describe cómo podemos utilizarla para mejorar mucho nuestro razonamiento.

Razonamos sobre causa y efecto, razonamos sobre obligación y permiso, razonamos sobre conflictos entre nuestras creencias y el mundo real. Sobre todo, razonamos cuando investigamos sobre lógica, ciencia y tecnología, y este libro considera tres casos de análisis: el uso de las analogías que hicieron los hermanos Wright al inventar el avión, el razonamiento probabilístico y deductivo de los criptoanalistas que rompieron el código Enigma alemán durante la Segunda Guerra Mundial y el razonamiento inductivo y explicativo del Dr. John Snow al descubrir la forma de transmisión del cólera, antes de que nadie conociera la existencia de las bacterias.

Este libro esboza una teoría de los procesos mentales para cada uno de los diferentes tipos de razonamiento. Los componentes de la teoría se han simulado mediante programas de ordenador, y sus predicciones se han valorado mediante experimentos psicológicos. Los experimentos han analizado cómo hacemos deducciones, resolvemos problemas, detectamos y conciliamos inconsistencias, creamos contraejemplos, inferimos probabilidades y llevamos a cabo la ingeniería inversa. Y mediante el escáner cerebral se ha empezado a identificar qué regiones del cerebro se activan al razonar.

He intentado escribir un libro que proporcione su propio sistema de soporte vital intelectual. El desarrollo de los temas en el texto principal está destinado a todos los lectores, el glosario es para aquellos que se inician en este campo y las notas al final del libro contienen información técnica y erudita destinada a los especialistas. Mi intención ante todo es explicar cómo razonamos, y hacerlo de

forma que satisfaga a todos los lectores que estén interesados en las respuestas científicas a las preguntas que planteaba al principio. Entre ellos se incluyen, espero, colegas del campo de las ciencias cognitivas, expertos de otros campos, estudiantes, adictos a los problemas lógicos y el lector común, como Virginia Woolf apodó una vez a cualquiera que lee por diversión.

CAPÍTULO 1

Introducción

¿Quiere que su mente sea inmortal? Su mente es quien usted es. Gobierna lo que piensa y lo que hace. En el pasado, cuando las personas morían dejaban tras de sí solo unos pocos recuerdos: cartas, fotografías, un puñado de cintas y vídeos. Pero, hoy poseemos la tecnología para preservar su mente después de la muerte, de modo que sus descendientes –incluso aquellos que nunca le conocieron– puedan interactuar con usted eternamente. Puedan hablar con usted, y puedan saber lo que usted cree y lo que usted piensa. Puedan incluso conocer sus nuevos pensamientos sobre cosas que no existían cuando usted vivía. Puedan seguir enseñándole.

¿Cómo se produce este milagro? Cada semana usted interactúa con un ordenador durante veinte minutos, contestando sencillas preguntas. El ordenador construye una representación de su conocimiento y sus creencias, y crea un programa que piensa de la misma forma en que usted lo hace. Veinte minutos por semana es una parte minúscula de su vida, pero al cabo de tres años, su mente está capturada en un ordenador. Ya puede seguir existiendo eternamente. Antes de que el Sol explote, los seres humanos dejarán el sistema solar. Se llevarán con ellos un amplio archivo de mentes humanas. Su mente podría estar entre ellas.

Este anuncio es ciencia ficción. Pero, las barreras para la construcción de este programa no son claras. Podría mostrar simplemente lo ignorantes que somos.

Los programas pueden tener acceso a enormes cantidades de información, y la cantidad de información en la Red es mucho mayor que la que cada uno de nosotros adquiere durante toda su vida. Nuestra ganancia neta de información es equivalente a más o menos mil novelas palabra por palabra (para encontrar referencias sobre las investigaciones relacionadas con el tema véanse las notas al final del libro)*. Los científicos cognitivos, sin embargo, tienen todavía que

diseñar un programa de ordenador que pueda pensar de la misma forma en que lo hacemos nosotros. Las teorías psicológicas son rudimentarias, y demasiado simples como para explicar el pensamiento que intervino en su propia construcción. Los robots perciben el mundo utilizando burdas imitaciones de nuestras capacidades perceptivas. La inteligencia (y la estupidez) humanas son tan difíciles de entender que muchos teóricos han abandonado la tarea en favor de la "inteligencia artificial". Diseñan programas que al menos resuelven problemas, si bien es cierto que por caminos muy alejados del pensamiento humano.

Podría pensar usted que el cerebro trabaja de una forma que un programa de ordenador no puede imitar. Sin embargo, cualquier descripción clara de un proceso, desde el tiempo atmosférico a la economía, puede simularse con un programa de ordenador. Esta conjetura se hizo hace setenta años y se llamó, partiendo de sus dos creadores, la tesis Church-Turing. La tesis no se puede probar, puesto que no existen test para valorar qué se considera una descripción clara. Pero, se refutaría con una demostración de que un conjunto de instrucciones comprensibles no se pueden implementar en un programa. Si las instrucciones para implementar su teoría de la mente fueran claras y usted pudiera probar que no se pueden simular en un programa de ordenador, habría usted refutado la tesis Church-Turing. La refutación sería más importante que su teoría. Ninguna dificultad para imitar la mente se refiere a los ordenadores y todas las dificultades se refieren a la no existencia de una teoría sobre cómo funciona esa mente.

Muchos filósofos han dudado si es posible tener una teoría científica del pensamiento –un escepticismo que se remonta al menos hasta Descartes–. Él pensaba que el libre albedrío condenaba al fracaso la tarea. Dostoyevsky se hacía eco de su opinión. Y, más recientemente, el novelista Arthur Koestler apostó con un psicólogo a que este no podría hacer una predicción acertada sobre el comportamiento de Koestler durante el mes siguiente. El psicólogo, Stuart Sutherland, debería haber ofrecido una apuesta de cien libras respecto a que Koestler no le daría mil libras. Pero Sutherland rechazó esta idea por injusta.

El quid de esta controversia es simple: los psicólogos no pueden predecir lo que pensará usted el viernes por la mañana o incluso cuál será su siguiente pensamiento, pero esta incapacidad no condena al fracaso una ciencia de la mente. Ni condena al fracaso un proyecto de simulación de su mente, porque usted no puede predecir cuál será su siguiente pensamiento. Puede hacer una

profecía autocumplida: “Pronostico que pensaré sobre el verde”, y a continuación puede intentar formar una imagen visual de una hoja verde. Pero puede ser difícil cumplir sus propias predicciones, p. ej.: “No voy a pensar en George W. Bush.” A veces nuestros pensamientos nos cogen por sorpresa. La mayoría de las veces simplemente tenemos pensamientos. La psicología es como la biología evolucionista. Los biólogos no pueden predecir qué especies serán las siguientes en evolucionar, pero eso no hace imposible una biología científica. La ciencia depende de las consecuencias observables de las teorías, y la biología da lugar a predicciones evaluables. La psicología también conduce a predicciones evaluables.

* * *

El pensamiento parece continuo, pero en el flujo de la consciencia solo se refleja su superficie. De ahí que para imitar el pensamiento los psicólogos necesiten considerar sus variedades. Algunas sencillas diferencias, que esquematizaré en forma de preguntas, dan lugar a una amplia taxonomía adecuada para nuestros propósitos.

¿Tiene un proceso de pensamiento una finalidad inmediata? Si no la tiene, los psicólogos lo llaman “asociativo”. Es decir, una idea lleva a otra en ausencia de cualquier diseño, plan o propósito. La persona que piensa tiene una ensoñación o está absorta. Las asociaciones tienen una larga historia en filosofía y psicología y podría pensar usted que son fáciles de entender. En realidad, ninguna teoría explica cómo hacemos asociaciones entre ideas. Es bastante fácil explicar cómo una palabra lleva a otra mediante una asociación. Usted dice, “pan”, y yo digo “mantequilla”, porque en nuestra mente las palabras se almacenan enlazadas unas con otras. En el pensamiento asociativo, sin embargo, una idea puede llevar a otra, y la cadena puede dar lugar a nuevas ideas que nunca antes habíamos tenido. Yo digo, “no solo de pan vive el hombre”, y usted dice, “no, pero es mejor que solo de mantequilla”, sin haber tenido antes ese pensamiento. No puede haber, en su memoria, un enlace que le lleve a esa idea porque es nueva para usted. No hay nadie que entienda cómo las asociaciones producen ideas novedosas. Afortunadamente, mi preocupación es el razonamiento, no la asociación.

Suponiendo que el pensamiento tiene un objetivo, ¿el proceso es determinista? Lo es cuando cada paso del proceso está determinado por el estado del sistema en ese momento. Los ordenadores son deterministas (aunque parezca lo contrario). Asimismo, cuando realizamos cálculos mentalmente, a menos que haya interrupciones, pensamos de una forma determinista. En cambio, los lógicos han imaginado máquinas que son no deterministas. El estado de la máquina en un momento dado no determina lo que sucederá a continuación. Puede restringir las posibilidades, pero no determinar una en exclusiva. Pensar se parece mucho a eso. Nuestros pensamientos no parece que se desplieguen como un mecanismo de relojería sin que exista en ningún punto posibilidad de elección. Si pudiéramos volver atrás en el tiempo y volver a empezar sin ningún recuerdo de nuestros primeros pensamientos, la segunda vez podríamos proceder de forma diferente. Una falta de determinismo tiene un valor práctico. Nos permite acercarnos a los problemas de diferentes formas, y evita que quedemos atascados en un bucle de pensamientos obsesivos. Nadie sabe si el pensamiento humano es no determinista, pero es una suposición que resulta útil. Permite la ignorancia de los psicólogos –no pueden predecir cuál será nuestro próximo pensamiento–. En teoría, esa suposición no tiene coste, porque un ordenador determinista puede simular otro que es no determinista.

El pensamiento que tiene un objetivo y es no determinista se puede dividir en dos categorías: creación y razonamiento. Ya hablaré de la creatividad, pero mi principal preocupación es el razonamiento o inferencia –utilizo los dos términos como sinónimos– y la definición de razonamiento con la que trabajo es:

Un conjunto de procesos que construyen y evalúan implicaciones entre conjuntos de proposiciones.

Esta definición hay que desmenuzarla, así que por favor tenga paciencia conmigo. Hay dos tipos de implicaciones que corresponden a los dos tipos principales de razonamiento: deducción e inducción. Una deducción –o inferencia válida– es la que da lugar a una conclusión que debe ser cierta en caso de que las premisas sean ciertas. Si las premisas no son ciertas la conclusión podría ser cierta, pero no existen garantías. Cualquier otro tipo de implicación es una inducción. Los libros de texto definen a menudo inducción como

“razonamiento desde lo particular a lo general”, oponiéndola a la deducción, que definen como “razonamiento desde lo general a lo particular”. Ninguna de las dos definiciones es totalmente correcta. Como verá usted, se pueden hacer deducciones partiendo de proposiciones particulares y llegando a proposiciones particulares, y se pueden hacer inducciones partiendo de proposiciones generales y llegando a proposiciones generales.

Una manera precisa de distinguir entre deducción e inducción es a partir de sus efectos sobre la información. Cuantas más posibilidades elimina una proposición, más información contiene. Por ejemplo, la proposición:

Hace sol y hace calor.

transmite más información que:

Hace sol.

que deja abierto el que haga calor o no. Las inferencias válidas producen conclusiones que no transmiten más información que las premisas. Podría usted preguntase por qué, en ese caso, nos merece la pena hacer deducciones. La razón es que una conclusión deductiva puede hacer explícita una proposición que solo estaba implícita en las premisas. He aquí, por ejemplo, una deducción válida que partiendo de premisas particulares llega a una conclusión particular:

La paciente tiene varicela o tiene sarampión.

No tiene varicela..

Por tanto, tiene sarampión.

La conclusión no añade información a las premisas, pero hace explícita una proposición que contiene la única posibilidad consistente con ellas.

En cambio, las inducciones, independientemente de lo verosímiles que sean, van más allá de la información contenida en las premisas. Sus conclusiones transmiten más información que ellas: excluyen posibilidades compatibles con las premisas. Sin embargo, pueden dar lugar a conclusiones verdaderas, a pesar de que no hay garantía de que lo hagan incluso siendo verdaderas las premisas. Los primeros aviadores, por ejemplo, hicieron todo tipo de inducciones. "Dennos un motor", dijo Sir Hiram Maxim, el inventor del fusil que lleva su nombre, "y muy pronto les daremos una máquina voladora satisfactoria". Sin embargo, su propio aparato dotado con dos motores de vapor revoloteó sobre su pista en 1894 y se estrelló. Por el contrario, Wilbur Wright no puso la prioridad en los motores, sino que infirió, en cambio, que lo que resultaba crucial era el control del aparato. Más adelante, cuando examine por qué los hermanos Wright tuvieron más éxito que sus rivales, volveré a referirme a cómo razonó. Pero, es imaginable que la inducción de Wilbur Wright –partiendo de proposiciones generales para llegar a una proposición general– podría haber sido errónea, y que motores ligeros y potentes habrían superado el problema del vuelo.

Ser inteligente es ser capaz de pensar, y ser capaz de pensar es ser capaz de razonar. El razonamiento es el centro del pensamiento humano, porque sin él seríamos irracionales. Pero, ¿somos racionales? Los filósofos discuten continuamente sobre esta cuestión. Me refiero en este caso, no a la opinión de que nuestro inconsciente bulle con cavilaciones irracionales, que puede hacerlo, sino a si somos capaces de razonar de forma racional. Cada una de las respuestas imaginables para esta pregunta tiene sus defensores. La opinión más optimista es que nunca cometemos errores al razonar: somos tan intachables como los ángeles. Una opinión intermedia es que podemos arreglárnoslas, a veces acertando y a veces equivocándonos. Una opinión insidiosa es que solo aparentamos razonar, pero de hecho llevamos a cabo un proceso totalmente distinto. La evolución humana nos ha adaptado, pongamos por caso, para enfrentarnos a nuestro entorno de una forma probabilística, y aplicamos los

mismos métodos incluso al razonamiento deductivo. Una doctrina más insidiosa es que no razonamos en absoluto, sino que hacemos conjeturas restringidas por las características superficiales de las premisas. Quizá la opinión más deprimente sea la de que no razonamos en absoluto a menos que se nos haya enseñado lógica. Pero, la lógica en sí misma no se podría haber inventado si nadie hubiera sido capaz de razonar antes de su invención. Se puede concluir sin temor a equivocarse que parte de nosotros razonamos durante parte del tiempo.

Sí nos diferenciamos en nuestra capacidad. Como han mostrado los experimentos, las diferencias entre un sujeto y otro no podrían ser mayores. Hay quien lo hace bien en casi todos los problemas inferenciales –un resultado que es inexplicable partiendo de la opinión de que nadie puede razonar–. Otros hacen poco más que conjeturar –un resultado que es inexplicable partiendo de la opinión de que somos intachables desde el punto de vista de la lógica–. Las inferencias en sí mismas difieren en dificultad. Algunas son juegos de niños. Por consiguiente, hasta los niños pequeños pueden hacerlas. Otras están fuera del alcance de todos nosotros con excepción de los más dotados. E incluso existen las que están fuera del alcance de cualquiera con excepción de los alienígenas sobrehumanos.

Estos resultados eliminan los extremos del espectro de opiniones. No somos ni ángeles ni insectos. De acuerdo con la teoría que voy a defender, somos racionales porque captamos la fuerza de los contraejemplos. Incluso sin haber estudiado lógica, nos damos cuenta de que una inferencia no es buena si podemos imaginar para ella un contraejemplo. A partir de la infancia, salvo quienes sufren un grave daño cerebral, todos somos capaces de algún pensamiento racional. Sin embargo, todos cometemos errores. Todos somos incapaces de hacer ciertas inferencias. Todos podemos mejorar nuestras capacidades –y más adelante hablaré sobre cómo podemos hacerlo–. Pero ninguno de nosotros pertenece a los coros angélicos. Somos racionales –en un cierto grado, como sabía Aristóteles– pero el razonamiento es un asunto que da vértigo.

* * *

Los psicólogos no han dicho exactamente qué es la inteligencia, pero una conjetura verosímil es que una parte importante de ella es la capacidad para razonar. Los que razonan bien en experimentos realizados en el laboratorio lo hacen bien en test de inteligencia. Si usted vive en Estados Unidos y quiere hacer estudios de postgrado, hace un examen nacional llamado Graduate Record Examination (GRE). Y si quiere estudiar Derecho, hace un examen llamado Law School Admission Test (LSAT). Estos dos exámenes tienen en común pruebas para valorar la capacidad para razonar. La Tabla 1.1 muestra un ítem típico del LSAT.

Tabla 1.1. Una pregunta de razonamiento lógico típica del LSAT

La mayor parte del café normal se hace con granos de la variedad arábica porque la gran mayoría de consumidores prefiere su sabor generalmente más fuerte al del café hecho con granos de la variedad robusta. Los bebedores de café que se pasan al descafeinado prefieren, sin embargo, por abrumadora mayoría el café hecho con la variedad robusta, que se diferencia de la variedad arábica en que a su sabor no le afecta tanto el proceso de descafeinado. Dependiendo del tipo de grano utilizado, el proceso de descafeinado reduce o elimina varias sustancias, la mayoría de las cuales no afectan al sabor, excepto una de las que contribuyen a enriquecer el sabor del café.

¿Cuál de las conclusiones siguientes se justifica mejor con arreglo a las afirmaciones anteriores?

(A) La cosecha anual mundial de café de la variedad arábica no es lo bastante

(B) Los granos de la variedad arábica contienen más cafeína por unidad de pes

(C) Los bebedores de café que toman café descafeinado son casi exclusivamer

(D) El proceso de descafeinado extrae una mayor cantidad de la sustancia que

(E) Hay bebedores de café que pasan de tomar café normal hecho con granos

Análisis

A medida que lee cada una de las opciones, es probable que piense que el texto parece irrelevante para (A), (B) y (E), que no implica (C), pero que probablemente implica (D). Un análisis más detallado muestra que el texto se resume en las siguientes proposiciones pertinentes para la inferencia:

Los granos de la variedad arábica tienen un sabor más fuerte que los de la variedad robusta. El sabor de los granos de la variedad robusta se ve menos afectado por el proceso de descafeinado.

Por consiguiente, la opción (D) se deduce del texto y es la elección correcta.

▪

¿Qué sentido tiene este test? El Consejo de Admisión en la Escuela de Derecho [Law School Admission Council (LSAC)] que realiza el LSAT dice: “Las preguntas de razonamiento lógico evalúan su capacidad para entender, analizar, criticar y completar una variedad de argumentos.” Y ¿por qué deberían seleccionar, en parte, las escuelas de derecho a los candidatos por su capacidad para razonar en esos aspectos? La respuesta a corto plazo es que los test predicen el rendimiento durante el primer curso en la facultad de Derecho. La respuesta a largo plazo es que los abogados necesitan razonar bien. No pueden acabar confundidos con embrollos similares al que aparece a continuación, procedente de una declaración real, en relación con el plagio de un programa de ordenador (el “código fuente”), de un testigo experto:

Abogado: Así que aunque usted dice que los resultados de nuestra experta están fatalmente viciados porque no examinó el código fuente, usted tampoco examinó

el código fuente, ¿no?

Testigo: No.

A: ¿Quiere usted decir que es incorrecto?

T: Sí.

A: ¿Así que usted sí examinó el código fuente?

T: Sí.

A: Entonces muestréme en su informe dónde examinó el código fuente.

T: Donde describí que examiné los puntos que se correspondían “en el contexto en que aparecen”.

A: Pero dice usted que con el fin de conocer la verdad sobre lo que hace un programa hay que examinar todo el código fuente, y usted estaba intentando saber la verdad sobre el programa, ¿correcto?

T: No.

A: O sea, ¿usted está diciendo que no tiene que examinar el código fuente para entender la verdad?

T: No.

A: Entonces, ¿qué está usted diciendo?

T: Que para entender la verdad de lo que un programa hace, hay que examinar el código fuente.

A: Entonces, ¿por qué contestó usted que no, previamente?

T: Porque lo que usted decía era falso.

A: ¿Qué le pregunté que fuera falso?

T: Cada una de las preguntas a las que contesté “no”.

Lo que hace equivocarse al abogado es una única palabra que a menudo es crítica en el razonamiento, “todo”. El testigo examinó el código fuente pero niega que sea necesario examinar todo el código fuente para entender un programa.

* * *

Algunos sujetos son expertos razonadores capaces de hazañas asombrosas. Como ejemplo de pericia en el razonamiento inductivo, aquí tenemos una secuencia de inferencias realizada por uno de los beduinos compañeros de viaje de Wilfred Thesiger en los desiertos del sur de Arabia en la década de 1940. Thesiger escribe:

Unos días después encontramos algunas huellas. Yo ni siquiera estaba seguro de que fueran de camellos, porque el viento las había vuelto muy borrosas. Sultan se volvió hacia un hombre de barba gris que era célebre como rastreador y le preguntó de quién eran las huellas, y el hombre se volvió hacia un lado y las siguió durante un corto trecho. Luego se bajó de su camello, examinó las huellas en el punto en que cruzaban un terreno duro, desmenuzó algunas boñigas de camello con los dedos y montó de nuevo para reunirse con nosotros. Sultan preguntó, “¿Quiénes eran?”, y el hombre respondió, “Eran Awamir. Son seis. Asaltaron a los Januba en la costa sur y les quitaron tres camellos. Vinieron de Shama y abrevaron en Mugshin. Pasaron por aquí hace diez días”. No habíamos visto ningún árabe desde hacía diecisiete días y no vimos a ninguno durante otros veintisiete. A nuestra vuelta nos encontramos con algunos Bait Kathir cerca de Jabal Qarra, y cuando intercambiamos noticias, nos dijeron que seis Awamir habían asaltado a los Januba, habían matado a tres y se habían llevado tres de sus camellos. Lo único que nosotros no sabíamos ya era que hubieran matado a nadie.

Otra forma más de pericia, un tanto más retorcida, se muestra en una secuencia de inferencias que hizo una paciente cuando se le preguntó por qué pensaba que podría haber contraído el SIDA por haber tocado una fotografía de periódico del actor de cine Rock Hudson, que había muerto a causa de la enfermedad. Ella explicó:

El fotógrafo tenía que haber estado cerca de Hudson, porque la fotografía era un "primer plano".

Por tanto, el fotógrafo podría haberse contaminado.

Por tanto, cuando reveló el negativo, podía haberlo contaminado.

El negativo estuvo en contacto con la fotografía impresa y por tanto podía haberla contaminado.

El hombre que hizo la fotografía del periódico utilizó la fotografía en la impresora del periódico, y por tanto podía haber pasado su contaminación a la impresora del periódico. La impresora podía haber pasado la contaminación a la imagen de todos los periódicos.

Por tanto, cuando yo toqué el periódico, yo también podría haberme contaminado.

La paciente sufría un trastorno "obsesivo-compulsivo". Como la mayoría de los razonadores expertos, puede pensar en más posibilidades que las obvias. Se da

cuenta de que su conclusión es improbable, sin embargo, no puede rechazarla, y por tanto se obsesiona con esa posibilidad.

Como ilustran estos ejemplos, los psicólogos tienden a pensar en el razonamiento como la tarea de sacar conclusiones o de evaluarlas. Pero, hay otros tipos de tarea inferencial. Cuando detectamos que nuestras creencias son inconsistentes, estamos razonando. Asimismo, para citar de nuevo al Consejo de Admisión en la Escuela de Derecho, la capacidad "para entender, analizar, criticar y completar una variedad de argumentos" exige en todos sus apartados el razonar. Si imaginamos una posibilidad que debilita un argumento en favor de una determinada conclusión, hemos razonado –como mínimo tuvimos que decidir que, dada la posibilidad que imaginábamos, las premisas ya no apoyan esa conclusión–. En el laboratorio se han estudiado pocas veces problemas de este tipo, pero en la vida real aparecen.

* * *

Muchos científicos cognitivos suponen que la lógica desempeña un papel central en el razonamiento y, por tanto, es necesario que le dedique unas palabras. Una inferencia válida, como señalé anteriormente, da lugar a una conclusión que debe ser verdadera si las premisas son verdaderas. Si las premisas no son verdaderas, la conclusión puede o no ser verdadera. De ahí que está garantizado que un argumento da lugar a una conclusión verdadera si se cumplen dos condiciones: premisas verdaderas e inferencia válida. El que estas dos condiciones son independientes entre sí fue la base de la lógica como disciplina intelectual, que en Occidente se remonta a Aristóteles. Él inventó la lógica –al menos dice que lo hizo– y la lógica atañe a las inferencias válidas.

La lógica moderna distingue entre forma y contenido. Para ilustrar el contraste consideremos de nuevo la inferencia basada en una disyunción (una premisa que contiene "o"):

La paciente tiene varicela o tiene sarampión.

No tiene varicela.

Por tanto, tiene sarampión.

Por una parte, la inferencia se puede demostrar en una lógica que tiene reglas de inferencia formales, incluyendo una regla que encaja con su forma lógica

A o B.

No A.

Por tanto, B.

Estas reglas actúan sobre las formas de las premisas –sobre los símbolos más que sobre su contenido– y construyen una lógica de pruebas. Por otra parte, la lógica tiene una interpretación. Si una disyunción A o B es verdadera, entonces al menos una de las dos proposiciones que la forman, A y B, es verdadera. Los lógicos establecen las posibilidades en una tabla, la llamada “tabla de verdad”. En el caso en que la premisa no A es verdadera, entonces A es falsa, y por tanto la única forma de que la premisa disyuntiva pueda ser verdadera es si B es verdadera. Por ello, la veracidad de ambas premisas implica que B debe ser verdadera: la inferencia es válida. Las tablas de verdad se refieren al contenido último de las proposiciones, su veracidad o falsedad, y construyen un sistema de interpretaciones.

Los lógicos han probado varias relaciones entre los dos sistemas de inferencia –

pruebas e interpretaciones–. Respecto a la lógica del “o”, “y” y “si” (la llamada lógica “booleana” o”proposicional”), han mostrado que toda inferencia que es válida utilizando tablas de verdad también se puede probar utilizando reglas formales de inferencia, y viceversa. Pero, su descubrimiento más sorprendente es que las equivalencias de este tipo fracasan en el caso de lógicas más potentes: las interpretaciones dan lugar a inferencias válidas que no se pueden probar con ningún sistema consistente de reglas formales. Este descubrimiento lleva de nuevo a una brecha insalvable entre forma y contenido.

La lógica es diferente del lenguaje natural. A la lógica no le afecta el tiempo: la interpretación de sus oraciones no depende de su contexto. Pero, en la vida real, la persona que habla utiliza una oración para establecer una proposición con una significación que casi siempre depende del contexto en que la pronuncia. La oración “yo puedo oírte ahora” (tan frecuente hoy en día en los lugares públicos) significa proposiciones diferentes dependiendo de su contexto. Usted entendió el significado de la oración cuando la leyó hace un momento, porque entiende el idioma. Cuando alguien la utiliza en la realidad, su conocimiento del contexto le capacita para entender su significación –la proposición concreta que expresa, que puede ser verdadera o falsa, dependiendo de a quién se refiere el “yo”, a quién se refiere el “te” y el momento en que se pronuncia–. Si “yo” se refiere a mí y “te” se refiere a mi mujer, y lo digo mientras escribo esta oración, entonces la proposición es verdadera, porque en este preciso momento yo puedo oír a mi mujer. Pero, si “yo” se refiere a mí y “te” se refiere a usted, el lector, entonces la proposición es falsa, porque ahora mismo yo no puedo oírle a usted. Por ello, una oración con un único significado puede expresar muchas proposiciones diferentes referidas a muchas situaciones diferentes. Ignoraré hasta donde sea posible esta distinción entre significado y significación.

En el pasado, muchos lógicos pensaron que las reglas formales de la lógica eran las leyes del pensamiento. Hilbert, uno de los mayores matemáticos del siglo veinte, escribió: “La idea fundamental de mi teoría de pruebas no es otra que describir la actividad de nuestra comprensión, establecer un protocolo de las reglas a las que realmente se ajusta nuestro pensamiento cuando avanza.” De hecho, incluso si el razonamiento humano fuera intachable, la lógica nunca podría ser una teoría psicológica completa, porque permite conclusiones válidas que nunca extraemos. Para ver por qué, necesito responder a una cuestión previa sobre el razonamiento: ¿qué es lo que hace el proceso? –o, para hablar de forma análoga a un programa de ordenador, ¿qué es lo que computa el proceso?

* * *

Cuando razonamos, podemos proponernos extraer conclusiones que sean válidas. Pero, hay muchas conclusiones válidas que nunca extraeríamos. Cuando los psicólogos defienden, como hacen a veces, que el razonamiento no es más que lógica, pasan por alto este punto. Para ver por qué, pregúntese a sí mismo lo que se deduce de estas dos proposiciones:

Ali fue el más grande de los boxeadores.

Hoy es miércoles.

La mayoría de nosotros dice: “nada”, porque las dos premisas no están relacionadas entre sí. No podíamos estar más equivocados, porque de dos premisas cualesquiera se deducen un infinito número de conclusiones válidas diferentes. Por ejemplo, las proposiciones anteriores dan lugar a esta serie de conclusiones:

Ali fue el más grande de los boxeadores y hoy es miércoles.

Ali fue el más grande de los boxeadores y hoy es miércoles y Ali fue el más grande de los boxeadores.

Ali fue el más grande de los boxeadores y hoy es miércoles y Ali fue el más grande de los boxeadores y hoy es miércoles.

Y así, ad infinitum.

Usted puede decir: Estas conclusiones son ridículas: pueden ser válidas, pero no nos dicen nada que no sepamos ya, y siguen repitiendo las mismas proposiciones una y otra vez. Ese es el asunto. Las conclusiones son ridículas; sin embargo, son válidas. Ninguna persona cuerda (que no fuera un lógico) haría estas inferencias. Por consiguiente, cuando extraemos conclusiones válidas, no extraemos cualquier conclusión por el hecho de que sea válida. Dadas las premisas:

Si la paciente tiene manchas con aspecto de granos, tiene varicela.

La paciente tiene manchas con aspecto de granos.

inferimos:

Tiene varicela.

Esta conclusión tiene en cuenta la información de ambas premisas, es parsimoniosa, y hace explícita una proposición que solo estaba implícita en ellas. Cuando un conjunto de premisas no da lugar a una conclusión de este tipo, como en el ejemplo sobre Ali y el miércoles, decimos que de ellas no se deduce nada. Y es verdad que de ellas no se deduce nada que tenga ningún interés o utilidad.

La conclusión de que la paciente tiene varicela no expresa toda la información de las premisas. No establece que la paciente tiene manchas con aspecto de granos. Pero no hay necesidad de repetir una premisa. El mantenimiento de

información implica que no hacemos inferencias, como por ejemplo:

La paciente tiene varicela.

Por tanto, la paciente tiene varicela o sarampión, o ambos.

La inferencia es válida, pero su conclusión desecha información: la premisa es consistente con una posibilidad en la que la paciente tiene varicela, mientras que la conclusión es consistente con tres posibilidades. Por tanto, sabemos más cuando conocemos la premisa que cuando conocemos la conclusión.

Cuando hacemos inducciones, estamos todavía más lejos de desechar información. La incrementamos. Nuestra intención es alcanzar conclusiones nuevas y parsimoniosas que sean verosímiles a partir de las premisas y de nuestro conocimiento. Nuestro razonamiento en la vida cotidiana no se produce en un vacío. (Las cosas pueden ser distintas en el laboratorio). Razonamos porque tenemos un objetivo más allá de extraer una conclusión. Tenemos un problema que resolver, una decisión que tomar, un plan que idear. El objetivo restringe los tipos de conclusión que extraemos. Dependiendo de nuestro interés en, pongamos por caso, un paciente determinado o una determinada enfermedad, extraeremos conclusiones diferentes partiendo de las mismas premisas sobre pacientes y enfermedades. El razonar, sin embargo, no es solo una cuestión de extraer conclusiones. También rechazamos algunas conclusiones. Podemos rechazarlas porque sabemos que son falsas o improbables. Pero también podemos rechazarlas porque están basadas en premisas falsas o en inferencias no válidas. Eso es, de forma esquemática, lo que hacemos cuando razonamos.

* * *

Lo que hacemos no es cómo lo hacemos. La diferencia entre ambas cosas es la

diferencia entre el jefe que encarga un programa de ordenador para hacer la nómina de la compañía y el programador que escribe el código para el programa. El primero establece qué debe hacer el programa; el segundo idea cómo lo hará. Muchos psicólogos teorizan solo sobre lo que la mente hace cuando razonamos. No hay nada malo en eso. Pero, cuando dicen, por ejemplo, que un tipo de razonamiento depende de procesos automáticos, implícitos y asociativos, usted debe saber que no hay ningún mecanismo asociado a esas etiquetas. Sería más exacto decir que el razonamiento se realiza mediante procesos desconocidos.

El cómo razonamos es un misterio que solamente con trabajo de despacho no se puede resolver. Exige también investigaciones experimentales. La dificultad del misterio será más fácil de captar si considero la variedad de posibles soluciones. Consideremos de nuevo esta inferencia:

Si la paciente tiene manchas con aspecto de granos tiene varicela.

La paciente tiene manchas con aspecto de granos.

¿Qué se deduce?

Es fácil inferir que la paciente tiene varicela –tan fácil que no somos conscientes de cómo lo hacemos–. Una posibilidad es que ya hayamos encontrado esa inferencia u otra similar a ella, y por tanto recuperamos nuestra conclusión anterior. Los teóricos se refieren a este procedimiento como razonamiento "basado en casos". Depende de una memoria para las inferencias, una forma de hacer corresponder las nuevas inferencias con las que hay en nuestra memoria –permitiendo algunas discrepancias en el contenido– y en la utilización de la conclusión que hay en la memoria para generar una conclusión para la inferencia actual. El rastreador de camellos podría estar utilizando un razonamiento basado en casos.

Ahora bien, tiene que haber una primera vez. Tenemos que hacer una inferencia para poder almacenarla en la memoria. Por definición, esta situación inicial no se puede resolver mediante el recuerdo de un caso anterior. ¿Podría estar equipada nuestra memoria con un amplio conjunto de inferencias innatas, de modo similar a como nuestros cuerpos están equipados con un amplio conjunto de anticuerpos innatos? Quizá; pero esta hipótesis traslada el problema de la explicación a una teoría del conocimiento innato. Lo que es más probable es que podamos razonar sin basarnos en recuerdos de inferencias previas.

Una teoría sobre cómo razonamos es que seguimos las “leyes del pensamiento”. Estas leyes se hacen explícitas en la lógica formal. No hay duda de que la gente puede adquirir estas leyes e incluso utilizarlas para resolver problemas difíciles. La pregunta que se plantea es si aquellos de nosotros que no tenemos una preparación en lógica formal seguimos las leyes del pensamiento. No somos conscientes de hacerlo. No las podemos describir. Pero aquí tenemos un esbozo de cómo funcionarían.

Un sistema formal de razonamiento es similar a un conjunto de reglas de lógica que actúan sobre las formas de las premisas. Una regla, por ejemplo, se aplica a oraciones con la forma lógica:

Si A entonces B.

A.

y da lugar a la conclusión:

B.

En la inferencia de la varicela que vimos previamente, las premisas tienen la

forma necesaria, donde A representa: “la paciente tiene manchas con aspecto de granos”, y B representa: “la paciente tiene varicela”. La regla arroja la conclusión: B, que se traduce de nuevo en la conclusión: La paciente tiene varicela.

Durante bastante más de un siglo, los teóricos han sostenido que las leyes del pensamiento son reglas formales de inferencia. El lógico George Boole lo sugirió a mediados del siglo diecinueve, pero su defensor más influyente fue Jean Piaget, un investigador suizo pionero en la psicología infantil. Piaget sostenía que la construcción mental de la lógica formal era el último gran paso en el desarrollo cognitivo de los niños, y que se producía en torno a los doce años. Otros teóricos han defendido reglas formales para el razonamiento inductivo. No obstante, diversas observaciones sugieren que el razonamiento no es un asunto formal.

Una de ellas es la aparición de errores sistemáticos al razonar. He aquí un ejemplo. Tenemos que seleccionar datos para averiguar si una afirmación general sobre cuatro cartas es verdadera o falsa:

Si una carta tiene una “A” en una cara entonces tiene un “2” en su otra cara.

Las cuatro cartas están extendidas ante nosotros: A, B, 2 y 3, y sabemos que cada una de ellas tiene una letra en una cara y un número en la otra. Una vez planteada esta tarea, tendemos a seleccionar la carta A y quizá la 2. No vemos la importancia de la carta 3. Pero, si esta carta tiene una A en su otra cara, la afirmación general es falsa. El difunto Peter Wason, que estableció el estudio experimental moderno del razonamiento, inventó esta tarea llamada “de selección”. Para cualquiera que haya estudiado lógica, el error de omisión en la tarea de selección es desconcertante. La conjunción de una A y 3 muestra que la afirmación inicial es falsa. El no seleccionar la carta 3 también está en contra de la afirmación de Piaget de que llegamos a dominar la lógica en la infancia. De ahí que, hace más de treinta y cinco años, Wason y yo iniciáramos investigaciones para intentar entender por qué se equivocaba la gente. Investigamos varias posibles fuentes de confusión, pero fuimos incapaces de hacer desaparecer el error de omisión. Wason sugirió entonces cambiar el

contenido de la afirmación general. Tal vez porque mi cabeza estaba llena de lógica formal, pensé que esta manipulación era insensata (aunque no lo dije), porque las reglas formales funcionan de la misma forma sea cual sea el contenido.

Wason siguió adelante con su experimento. Utilizó una afirmación general sobre los viajes:

Todas las veces que voy a Manchester voy en tren.

Las oraciones no siempre dejan traslucir su forma lógica, y, a pesar de las diferencias gramaticales, esta afirmación tiene la misma forma lógica que la afirmación previa referida a las letras y los números. Sin embargo, solo para estar seguro, Wason incluyó en el experimento la afirmación siguiente: Toda carta con una “A” en una cara tiene un “2” en su otra cara. Los participantes se equivocaron en este caso, pero se dieron cuenta de que un viaje en coche tenía relación con la veracidad o falsedad de la afirmación sobre los viajes. Si el destino del viaje en coche era Manchester, la afirmación era falsa. Un cambio que afectaba únicamente al contenido tuvo un efecto sorprendente sobre el razonamiento, aun cuando los dos tipos de contenidos tenían la misma forma lógica.

Estos resultados resultaban embarazosos para la teoría formal. Por una parte, el error sistemático de omisión con las letras y los números era contrario a las leyes del pensamiento plasmadas en las reglas formales. Por otra parte, un cambio que afectaba únicamente al contenido no debería tener efectos sobre la forma de actuar, dado que por definición las reglas formales no tienen en cuenta el contenido. Una posibilidad es, sin embargo, que el test de selección no consiga aprovechar nuestra capacidad deductiva. Algunos psicólogos tienen esa opinión. Yo no; pero no se ha llegado a un acuerdo sobre el tema que satisfaga a todos.

Una teoría alternativa del razonamiento, que se desarrolló inicialmente en el campo de la inteligencia artificial, desplaza el método general hacia un manejo más sensible del contenido. Las reglas formales afectan a términos lógicos, como “no” y “si”. Pero las reglas formales se pueden formular de manera que encajen también con otros tipos de palabras, incluyendo nombres y verbos

concretos. Por ejemplo, un sistema creado específicamente para razonar sobre pacientes y enfermedades podría mantener registros, como por ejemplo:

Paciente: Mary Smith Síntomas: manchas con aspecto de granos.

Podría contener reglas para hacer inferencias, como por ejemplo:

Si el objetivo es mostrar que el paciente x tiene varicela

y el paciente x tiene manchas con aspecto de granos

entonces el objetivo está conseguido.

Reglas de este tipo son específicas del contenido, porque son sensibles a ciertos contenidos, como pacientes, manchas y varicela. Se utilizan en programas de ordenador que plasman la pericia humana y aportan asesoramiento en temas como el diagnóstico médico. Las reglas pueden incluir contenidos detallados, p. ej.:

Si el lugar de cultivo es la sangre, el organismo es gram-negativo, la morfología del organismo es bastón, el escozor del paciente es grave, entonces hay algunos indicios (0.4) de que la identidad del organismo es pseudomonas.

A pesar de su nombre, las reglas específicas del contenido funcionan de una manera formal. El procedimiento sigue siendo asociar la forma de las premisas a

reglas, pero la noción de forma se ha ampliado para incluir palabras que no son términos lógicos. Un programa que utilice la regla anterior no sabe nada sobre el significado de la palabra “sangre”, o la significación de las proposiciones sobre pacientes, pero puede hacer inferencias formales.

Ningún programa actual capta la significación de una oración que escribimos en el teclado del ordenador. No se enfrentan con la significación real más de lo que los programas que simulan huracanes se enfrentan a vientos reales. Hasta que los sistemas computacionales –los robots, tal vez– puedan percibir el mundo, acceder a sus propios estados internos y actuar sobre el mundo, solo pueden simular que captan la significación. Nosotros, los seres humanos, somos las únicas entidades conocidas que pueden captar la significación de las proposiciones, e imaginar las situaciones que describen.

Consideremos una vez más estas premisas:

Si la paciente tiene manchas con aspecto de granos entonces tiene varicela.

La paciente tiene manchas con aspecto de granos.

Cuando entendemos la primera premisa, nos damos cuenta de que es compatible con al menos dos posibilidades. Una posibilidad es que la paciente tenga manchas con aspecto de granos y varicela, y la otra posibilidad es que la paciente no tenga manchas con aspecto de granos. Pero también captamos que la segunda premisa excluye esta última posibilidad. Solo resta la primera posibilidad, y por tanto se deduce que la paciente tiene varicela. Una inferencia es válida si su conclusión se cumple en todas las posibilidades en que se cumplen las premisas, y por tanto garantiza que su conclusión es verdadera en caso de que sus premisas sean verdaderas. No es válida si hay un contraejemplo –una posibilidad consistente con las premisas pero no con la conclusión–. No existe ningún contraejemplo para la conclusión anterior, y por ello la inferencia es válida. Hemos hecho una deducción utilizando, no la forma lógica de las premisas, sino sus contenidos, para imaginar posibilidades.

En el ejemplo anterior describí las posibilidades, pero, ¿cómo se representa en nuestra mente la significación de las oraciones? Estas son las pistas para una solución:

Es posible que a partir del significado de las oraciones en un discurso conexo el oyente establezca implícitamente un modelo de la narración más abreviado y no especialmente lingüístico... En los puntos en que el modelo está incompleto, incluso puede inventarse material de manera inconsciente para hacer que el recuerdo esté más lleno de contenido o sea más verosímil –un proceso que tiene su paralelismo en la construcción inicial del modelo–. Un buen escritor o narrador quizá tiene la capacidad de iniciar un proceso muy similar al que se produce cuando realmente estamos percibiendo (o imaginando) hechos en lugar de estar solamente leyendo u oyendo algo sobre ellos.

La investigación posterior ha apoyado esta suposición. Construimos modelos mentales de las situaciones, y utilizamos esos modelos para representar posibilidades.

Durante la última década, la investigación sobre modelos mentales y razonamiento ha florecido tanto que nadie puede seguirle la pista al completo. Y la teoría en sí misma ha sufrido una metamorfosis. La distinción entre lógica y razonamiento, que se presagiaba en los descubrimientos de Wason, ha alcanzado actualmente un grado mucho mayor del que nadie podría haber imaginado. La lógica es una herramienta esencial para todas las ciencias, pero no es una teoría psicológica del razonamiento. El ámbito de la teoría de modelos se ha ampliado también para incluir todos los principales tipos de razonamiento, y dar una explicación exhaustiva de ellos –desde los procesos inconscientes de bajo nivel hasta el desarrollo de estrategias inferenciales.

* * *

Memoria, reglas formales, reglas específicas de contenido y modelos de

posibilidades. ¿Hay alguna otra forma en que podamos razonar? Quizá, pero es improbable. A veces mis colegas preguntan: ¿Y los sistemas que imitan las operaciones de una red idealizada de neuronas mientras están aprendiendo a realizar una tarea? Las llamadas redes “conexionistas” pueden hacer inferencias rudimentarias. Pero, para hacer inferencias con el modesto grado de complejidad del que nosotros somos capaces, puede que estas redes tengan que implementar una de las teorías que ya he descrito.

¿Cómo implementan las neuronas reales del cerebro nuestros procesos de razonamiento y en qué regiones del cerebro están localizadas? Estas preguntas se han empezado a investigar, y hablaré de ellas. Pero se sabe poco sobre las respuestas. El cerebro no puede ser en absoluto una herramienta computacional, aunque esto no implicaría que no se pueda simular en un programa de ordenador. Puede que no utilice en absoluto representaciones mentales, como el filósofo alemán Heidegger y sus herederos intelectuales han sostenido. Puede que no utilice mecanismos mentales inconscientes, como han sostenido otros filósofos. Sin embargo, estas opiniones no pueden dar cuenta de nuestra capacidad para razonar. Razonar es llevar a cabo un proceso mental. La introspección no nos revela cómo funciona ese proceso. (Si lo hiciera, los psicólogos habrían entendido cómo funcionaba hace mucho tiempo.) En consecuencia, el proceso es inconsciente. Cuando razonamos sobre algún suceso que hemos recordado, ese suceso debe estar representado en nuestra mente. Eso parece evidente, aunque no para los que son escépticos. Discutir con ellos es una extraña experiencia –como yo me imagino que sería el discutir con alienígenas de Alfa Centauri–, no obstante, no tienen una teoría sobre cómo razonamos. Suponiendo que nuestro razonamiento dependa de al menos uno de los cuatro tipos de procedimiento que he descrito, ¿cuál es el candidato más verosímil? ¿Y cuáles son los detalles de su implementación? A eso es a lo que está dedicado el resto del libro.

Notas al pie

* Esta es la única nota a pie de página del libro.

Parte I

El mundo en la mente consciente

Una oración pronunciada hace surgir un mundo

En el que todo sucede como ella dice;

Dudamos de quien habla, no de la lengua que oímos:

Las palabras no tienen ninguna palabra para las palabras que no son ciertas.

Auden

CAPÍTULO 2

Iconos e imágenes

¿Está usted cómodamente sentado? Si es así, responda por favor a esta pregunta:

¿En qué dirección está mirando: Norte, Sur, Este u Oeste?

Si se parece usted a mí, no conocerá la respuesta exacta. Todo lo que sé es que estoy mirando más o menos en dirección al Oeste. No me oriento según los puntos cardinales, pero no lo necesito. Los Guugu Yimithirr, una tribu australiana, siempre están orientados según su equivalente a nuestros puntos cardinales. Steve Levinson, un lingüista y antropólogo, filmó a un grupo de hombres charlando, y lo que resultaba sorprendente era la frecuencia con que señalaban en esta o aquella dirección. Llevó a algunos de ellos a un zigzagueante viaje en autobús por el interior despoblado de Australia, y detuvo el autobús cada cierto tiempo. Les pidió entonces que señalaran en la dirección de diversos puntos que les eran familiares, como por ejemplo su casa. Aquellos hombres podían realizar esa tarea mucho mejor de lo que lo haríamos nosotros, quizá porque sus ancestros no se podían permitir perderse en esa zona. En una ocasión se produjo un fallo importante y todavía se ríen con ello. Uno de sus compañeros se subió a un avión, y durante el viaje perdió la orientación –un incidente que le causó ansiedad hasta que fue capaz de reorientarse cuando estuvo de nuevo en tierra.

Incluso sin ser capaz de saber su orientación según la brújula, usted debería ser capaz de contestar a esta pregunta:

¿En qué dirección está mirando en relación a la puerta principal del edificio

dentro del que se encuentra sentado?

Si ahora mismo no está en ningún edificio, imagine que está en uno que le sea familiar. La pregunta es fácil a menos que usted esté en un edificio muy grande.

Una pregunta más. Imagine que está tumbado en la cama en su casa, ¿cuál es el camino desde allí hasta la puerta principal? Trácelo en el aire con el dedo.

Esta tarea tendría que haber sido bastante fácil. Utilizó usted su mano dominante. Y movió el dedo índice a lo largo del camino desde la cama a la puerta del dormitorio, bajando las escaleras si las hay y a través de la casa hasta la puerta principal. Puede que haya tenido imágenes visuales de partes del recorrido, o incluso alguna imagen mental cinestésica mientras bajaba las escaleras. (La imagen mental cinestésica representa movimientos de las extremidades y otras partes del cuerpo.) Sin embargo, si se parece usted a mí, no habrá sido consciente de ninguna imagen, sino que simplemente habrá movido el dedo a lo largo de un camino tridimensional correcto.

¿Cómo somos capaces de hacer estas tareas? La respuesta amplia es que nos basamos en una representación mental del mundo cuyos contenidos están a disposición de la consciencia. Yo sabía, por ejemplo, que estaba mirando más o menos hacia el Oeste, porque sabía que estaba mirando más o menos en perpendicular a Nassau Street, la calle principal de Princeton, que va aproximadamente –muy aproximadamente de Norte a Sur–. Así que, aunque no tenía imágenes visuales, supe inmediatamente que estaba en perpendicular a Nassau Street. Tenía una representación interna de esta relación. Su contenido me capacitaba para inferir la orientación.

Una de las mayores hazañas de la orientación mental es la capacidad para la navegación de ciertos habitantes de las Islas Carolina en el Océano Pacífico. Estos navegantes pueden manejar sus canoas con balancín a lo largo de distancias de varios cientos de millas a través del mar abierto y llegar con éxito a otra pequeña isla. Lo hacen sin brújula ni ningún otro instrumento de navegación. Tienen un escaso margen de error: un grado o dos de desviación, y no acertarán con su isla de destino y como resultado morirán. Los navegantes conocen las posiciones de todas las islas, y los vientos y corrientes marinas predominantes. Pero, lo más importante de todo, conocen la salida y posición de

las principales estrellas. Imaginan que el bote está quieto y que las islas son arrastradas a su lado por la corriente, y la estrellas giran sobre sus cabezas a medida que la Tierra gira. Como punto de referencia, mantienen en mente una isla que está fuera del borde de su ruta y muy por debajo del horizonte. Esta maravillosa representación les permite navegar mediante cálculos estimados, donde “estimado” representa “deducido”[1].

* * *

Cuando los psicólogos mostraron por primera vez que los animales parecían construir “mapas” cognitivos de los laberintos por los que corrían, los escépticos sostuvieron que quizá los animales elaboraban un registro de su secuencia de respuestas al correr por un laberinto. Estos “conductistas” llamaron a cada componente de esta secuencia una “respuesta fraccionaria anticipadora de meta”, eso era lo más que concederían en cuanto a que los animales, incluidos nosotros, tuvieran una vida mental. Según esta explicación, mi mente no contiene una representación de las relaciones espaciales entre mi silla y Nassau Street. Todo lo que contiene es una secuencia de respuestas “fraccionarias”: girar a la izquierda, avanzar diez pies, pasar a través de la puerta del salón, girar a la izquierda, salir por la puerta principal, etc. Fue fácil mostrar, por el contrario, que las ratas construyen mapas internos. Aprendían una ruta tortuosa para llegar a una caja que contenía comida. Cuando el experimentador bloqueaba esta ruta, eran capaces de elegir un camino directo a la caja de comida. No podrían haberlo hecho si solo hubieran tenido un registro de respuestas. Estudios sobre la búsqueda de alimento en animales también refutan a los escépticos: los chimpancés y los pájaros pueden recobrar comida de diversos escondites en un orden diferente de aquel en que se les mostraron los escondites.

La orientación espacial no es exclusiva de los mamíferos. Las abejas obreras que encuentran néctar realizan una danza en la vertical del panal cuando vuelven a la colmena. La maniobra principal en esta danza es hacer oscilar sus abdómenes a medida que suben y bajan por el panal con un ángulo determinado en relación con la vertical. El ángulo indica con qué ángulo, en relación al sol, deben volar las otras abejas para llegar a la fuente del néctar. La velocidad de las oscilaciones indica su distancia. Las abejas comunican un mensaje simple, y su sistema de

navegación mediante cálculo “estimado” las lleva al néctar. También les permite hacer inferencias espaciales: después de haber volado hasta el néctar por un camino tortuoso, su danza señala la ruta directa.

Quienes proponen las respuestas fraccionarias tienen algunas dificultades para explicar lo que hacen los pájaros y las abejas. Los Guugu Yimithirr hacen aún más difícil esta explicación, porque ellos no han aprendido una secuencia de respuestas que les pueda capacitar para señalar en la dirección de una ciudad lejana. Deben tener una representación interna de la geometría de su entorno, que actualizan con cada serpenteo y giro del camino. Usted y yo también podemos poner trabas a la hipótesis de las respuestas fraccionarias. Podemos trazar en el aire un camino por el que podríamos volar como Peter Pan desde nuestra cama, salir por la ventana, descender a la planta baja y a través de paredes macizas llegar al fogón en la cocina. Nunca hemos tomado ese camino, y por tanto no hemos adquirido su secuencia de respuestas. Lo que nos permite imaginar esa ruta es que tenemos una representación mental de la distribución espacial de nuestras casas. Incluso sujetos que son ciegos construyen estas representaciones a base de recorrer su casa.

El hipocampo es el órgano del cerebro que parece ser crucial para la orientación espacial. Reside en el fondo del lóbulo temporal con muchas conexiones con el córtex de la superficie. También es crucial para establecer nuevos recuerdos. Un ejemplo trágico se refiere a un paciente al que se le habían extirpado partes del hipocampo. Era incapaz de recordar nada nuevo sucedido después de la operación. Todos los días eran el inicio del resto de su vida, porque no podía recordar nada de lo que había sucedido el día anterior. Estas diversas funciones del hipocampo parecen tener en común el procesamiento de relaciones espaciales y de otro tipo.

* * *

Supongamos que en lugar de examinar una habitación leemos una descripción, como por ejemplo:

Es una habitación del primer piso, que da a la calle, destinada originariamente a ser la sala. Las puertas dobles están en el centro de la pared del fondo; y las personas que entran por ellas encuentran, a su derecha, dos altos archivos, formando ángulo recto entre sí, adosados contra las paredes. En ese rincón hay una mesa de escribir y sobre ella un fonógrafo, un laringoscopio... [y aquí el autor enumera otra parafernalia].

Más adentro, del mismo lado, hay una chimenea, con una confortable butaca forrada de cuero en el costado del hogar que está más cerca de la puerta, y un cubo para el carbón. Sobre la repisa de la chimenea hay un reloj. Entre la chimenea y la mesa del fonógrafo hay un revistero.

Al otro lado de la puerta central, a la izquierda del visitante, se encuentra un armario de cajones poco profundos... El rincón, más allá, y la mayor parte de la pared lateral, están ocupados por un piano de cola, con el teclado situado hacia el lado más alejado de la puerta, y un banco que se extiende a lo largo de todo el teclado...

La descripción forma parte de las orientaciones de Bernard Shaw respecto a la distribución del estudio del profesor Higgins en su obra Pygmalion (también conocida como My Fair Lady en su transfiguración en musical debida a Lerner y Loewe). La descripción nos resulta lo bastante clara como para elaborar un plano aproximado de la habitación. La Figura 2.1 muestra las estructuras en alambre de los objetos de la habitación realizadas con un programa de ordenador para construir modelos tridimensionales (Design Workshop_Lite). El programa permite al usuario ver la habitación desde cualquier punto de vista, e incluso caminar por ella. La Figura 2.2 muestra una vista de la habitación con los objetos en color, y la Figura 2.3 muestra otra. También nosotros, si oímos a alguien leer la descripción, podríamos imaginar la habitación, e incluso verla desde diferentes ángulos. Pero, nuestra representación no es tan completa como el modelo del ordenador. A medida que avanza la descripción, no podemos mantener presentes todos los detalles. No obstante, nuestra representación mental se parece a la del programa en que describe las relaciones espaciales entre los principales objetos del escenario, y la información “métrica” aproximada respecto a las distancias entre ellos. Esta afirmación se confirma con

pruebas psicológicas. Podemos hacer inferencias a partir de descripciones, p. ej., los tres objetos siguientes están alineados a lo largo de una pared en este orden: el reloj en la repisa de la chimenea sobre el fuego, el revistero y los archivadores. Más adelante volveré a hablar de cómo hacemos estas inferencias espaciales.

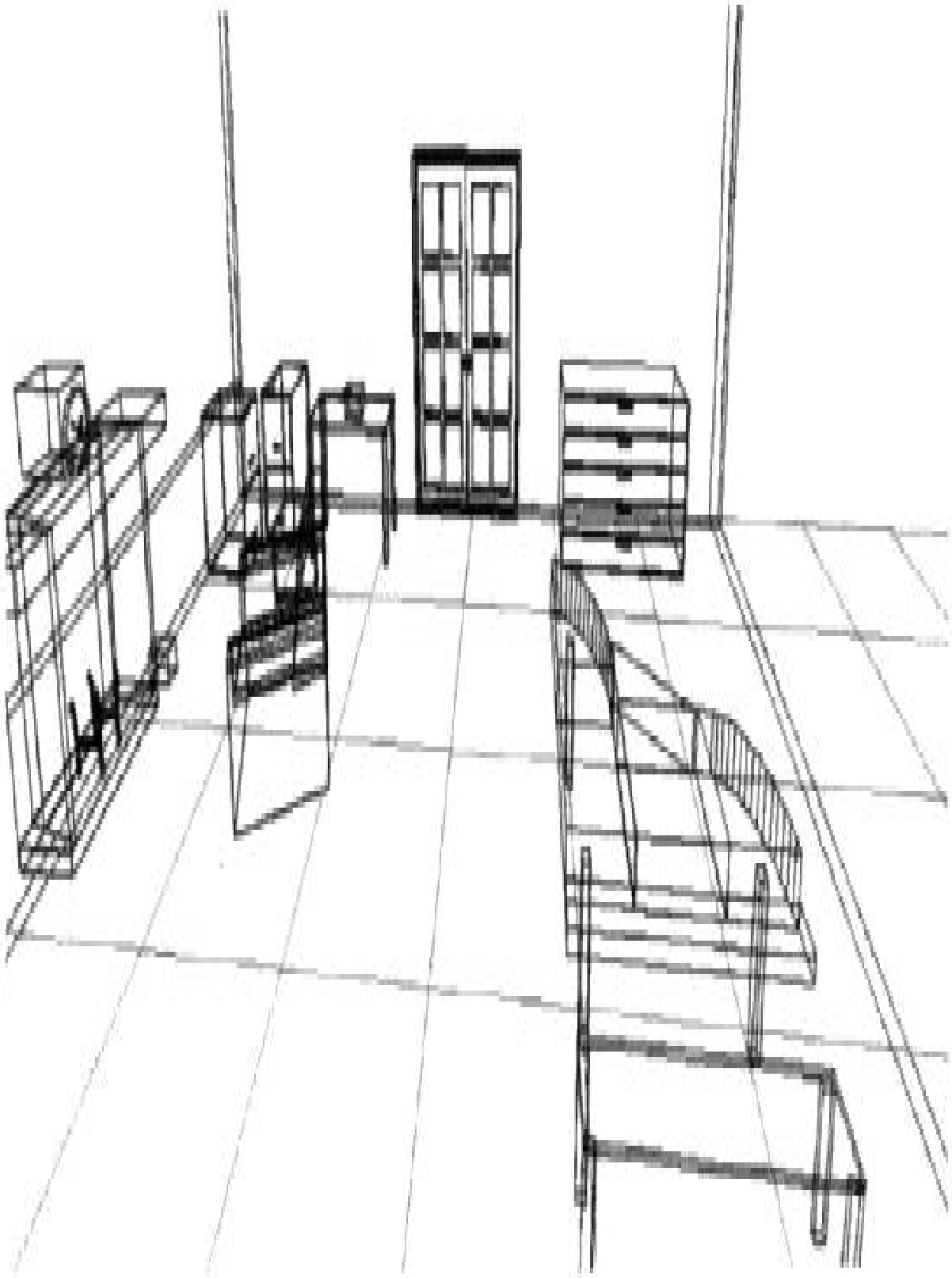


Fig. 2.1 El estudio del profesor Higgins (tal como se describe en Pygmalion de Bernard Shaw) en un esquema de los objetos hecho con el programa DesignWorkshop_Lite. La descripción es ambigua y por lo tanto la mesa podría estar situada entre los dos archivadores.

Fig. 2.2. Una perspectiva del estudio del profesor Higgins.

Fig. 2.3. Otra perspectiva del estudio del profesor Higgins.

Sin más preámbulos, asumiré que los seres humanos pueden construir representaciones internas de distribuciones espaciales, y me referiré a estas representaciones como modelos mentales de distribuciones espaciales. Un modelo mental es icónico, es decir, sus partes y las relaciones entre ellas se corresponden con las partes de la distribución y las relaciones entre ellas. En contraposición, la forma lógica de una oración no es icónica, porque su estructura está alejada de la estructura de la situación que describe. Un modelo espacial, sin embargo, es un icono de lo que representa.

* * *

¿Podría un modelo mental no ser nada más que una imagen visual? Para el Renacentista Leonardo da Vinci visualizar era pensar: los dos procesos eran uno y eran el mismo. El ojo era el instrumento del pensamiento, y la habilidad del artista para hacer cuadros era un medio mediante el cual realizar experimentos imaginarios. A finales del siglo diecinueve, el físico Ludwig Boltzmann escribió: “Todas nuestras ideas y conceptos son solo cuadros internos.” Escribió también: “La tarea de la teoría consiste en construir una imagen del mundo externo que solo existe internamente y debe ser nuestra estrella guía en el pensamiento y la experimentación; o sea, en completar, por así decirlo, el proceso de pensar y llevar a cabo globalmente lo que a pequeña escala ocurre dentro de nosotros cada vez que formamos una idea.”

Muchos expertos han señalado que la imagen mental visual es fundamental en las artes, las ciencias y la tecnología. Las anécdotas refuerzan esta fama. Friedrich August von Kekulé, por ejemplo, describió cómo en 1856 descubrió el anillo de la estructura química del benceno a partir de un sueño. Escribió:

Volví la silla hacia el fuego y di una cabezada. Los átomos estaban de nuevo retozando ante mis ojos. Esta vez los grupos más pequeños se mantenían modestamente al fondo. Mi ojo mental, que se había vuelto más agudo a base de las repetidas visiones de este tipo, podía ahora distinguir estructuras más grandes, organizadas como un colector; largas filas, a veces más estrechamente unidas; todas entretejiéndose y retorciéndose con un movimiento similar al de las serpientes. ¡Pero fíjate! ¿Qué era eso? Una de las serpientes había cogido su propia cola, y la forma daba vueltas burlonamente ante mis ojos. Me desperté como un relámpago.

La serpiente mordiéndose la cola, afirmaba, fue la pista crucial para la estructura circular del benceno.

En una de las primeras investigaciones sobre las imágenes visuales, Francis Galton, el primo de Darwin, llevó a cabo un estudio mediante cuestionarios entre sus compañeros científicos. El estudio reveló –para su consternación– que muchos de ellos afirmaban pensar sin utilizar imágenes. Sin embargo, muchos otros científicos han destacado el papel de la imagen visual en su pensamiento. El mismo Darwin escribió a Galton: “Recuerdo vívidamente los rostros de las personas que he conocido con anterioridad, y puedo conseguir que hagan cualquier cosa que yo quiera”. Quizá el mejor reconocimiento se produjo mucho después en una carta que Einstein escribió al gran matemático francés Hadamard:

Las palabras del lenguaje, tal como se dicen o escriben, no parecen desempeñar ningún papel en mi sistema de pensamiento. Las entidades físicas que parecen servir como elementos en el pensamiento son ciertos signos e imágenes más o menos claras que se pueden reproducir y combinar “voluntariamente”.

Hay, por supuesto, una cierta conexión entre esos elementos y los conceptos lógicos relevantes. También está claro que el deseo de llegar finalmente a conceptos conectados lógicamente es la base emocional del juego bastante impreciso con los elementos antes mencionados. Pero considerado desde un punto de vista psicológico, este juego combinatorio parece ser el rasgo esencial

en el pensamiento productivo –antes de que se produzca cualquier conexión con la construcción lógica mediante palabras u otro tipo de signos que puedan ser comunicados a otros.

Fundamentar una teoría psicológica en introspecciones, no importa lo distinguido que sea su origen, es construir sin bases. Es imposible establecer la verosimilitud de informes subjetivos. En el peor de los casos, pueden ser un fraude, como han afirmado algunos respecto al sueño de Kekulé; en el mejor, pueden estar equivocados, porque ninguno de nosotros tiene acceso a los manantiales del pensamiento.

* * *

Para averiguar lo buenas que son sus imágenes mentales intente resolver este problema:

Tengo miles y miles de agujas muy finas, que sostengo en un montón entre las manos. Las lanzo hacia arriba en el aire, aplicando a cada una fuerza aleatoria. Caen al suelo, pero, antes de que ninguna de ellas lo toque, las paro mágicamente en el aire. Muchas agujas están en posición horizontal o casi horizontal, y muchas están en posición vertical o casi vertical. ¿Es probable que haya más agujas de la primera categoría, más de la segunda categoría o habrá aproximadamente las mismas de ambas categorías?

La mayoría de la personas a las que he planteado este problema dicen que habrá aproximadamente las mismas de las dos categorías. Están equivocadas. Piénselo de esta forma: una aguja está horizontal si apunta al norte, sur, este u oeste o a cualquier orientación intermedia. No hay equivalente para una aguja vertical: debe apuntar hacia arriba con un ángulo de más o menos 90º con la horizontal. De ahí que la respuesta correcta es que lo más probable es que haya más agujas

horizontales.

He aquí otro problema similar:

Tengo miles y miles de discos circulares muy finos, que sostengo en un montón entre las manos. Los lanzo hacia arriba en el aire, aplicando a cada uno una fuerza aleatoria. Caen al suelo, pero, antes de que ninguno de ellos lo toque, los paro mágicamente en el aire. Muchos discos están horizontales o casi horizontales, y muchos están verticales o casi verticales. ¿Es probable que haya más discos de la primera categoría, más de la segunda categoría o habrá aproximadamente los mismos de ambas categorías?

Ahora que conoce usted la respuesta al primer problema, puede conseguir resolver bien este. Pero, las personas que no han oído hablar de ese primer problema tienden a elegir cantidades aproximadamente iguales de ambos tipos de discos. Están equivocadas. Un disco vertical puede estar alineado de norte a sur, de este a oeste o en cualquier otra orientación intermedia. Pero no hay equivalente para un disco horizontal. Podría imaginarse una muesca en el borde del disco, y pensar que el disco se puede girar en el plano horizontal de manera que la muesca esté alineada con el norte, sur, este, oeste o cualquier orientación intermedia . Pero, si el disco está vertical, puede rotarse de manera similar de modo que la muesca esté alineada con las doce, la una, las dos, etc., dando la vuelta entera al reloj. Además, puede estar alineado con cualquier orientación de la brújula, independientemente de la posición de la muesca. De ahí que el disco vertical tenga más posibles posiciones. Estos dos problemas son difíciles porque somos malos imaginando objetos en tres dimensiones de forma independiente de nuestro punto de vista. Como en los intentos de visualizar el estudio de Higgins, estamos abrumados por la cantidad de información que tenemos que tener presente.

No todo se puede visualizar. Por ejemplo, se puede pensar en la mecánica cuántica. Una sutil ironía es que uno de sus principales expertos, el difunto Richard Feynman, tenía una notable capacidad para visualizar. Cuando la gente se queja de que es imposible visualizar cómo se producen los sucesos cuánticos, olvidan que muchos conceptos de la vida diaria son también imposibles de

visualizar. Auden se preguntaba cómo podríamos generar una imagen de “costumbre”. No podemos reducirla a características perceptibles, ni tampoco otros conceptos abstractos como posibilidad o propiedad. Podemos percibir diversos documentos o sucesos que sirven como prueba de que yo soy propietario de un coche, pero no podemos percibir la relación de propiedad en sí misma.

Quizá tenemos dos sistemas independientes de razonamiento. Un sistema trabaja con imágenes de entidades concretas y el otro sistema trabaja con representaciones abstractas. No somos conscientes de ningún tirón en el cambio de marchas intelectual cuando razonamos sobre una mezcla de contenidos abstractos y concretos. Otra posibilidad es que tengamos un único sistema de razonamiento que se basa en imágenes mentales de diversos tipos – visuales, auditivas, cinestésicas– y de alguna forma representemos las abstracciones en imágenes. Parece factible, pero parece más probable una tercera posibilidad: el cerebro representa todo mediante modelos.

Los modelos mentales no se deben confundir con imágenes visuales. Una imagen visual representa el aspecto que algo tiene desde un punto de vista determinado. Es una representación egocéntrica. En cambio, un modelo mental es una estructura abstracta similar a las disposiciones espaciales tridimensionales utilizadas en el programa de ordenador que simulaba el estudio del profesor Higgins. No es egocéntrica, sino independiente del punto de vista. El programa utiliza esta disposición tridimensional para proyectar un dibujo bidimensional del objeto sobre la pantalla del ordenador. De manera similar, nuestros cerebros podrían utilizar un modelo mental tridimensional subyacente de un objeto para construir una imagen bidimensional de su apariencia desde un determinado punto de vista. Puede uno preguntarse qué pruebas existen para apoyar esta afirmación. Describiré tres experimentos que la corroboran, pero antes quiero decir algo sobre los experimentos psicológicos.

* * *

Los experimentos con participantes humanos son interacciones sociales en un microcosmos, y por eso son sensibles a muchos factores imprevistos. Los

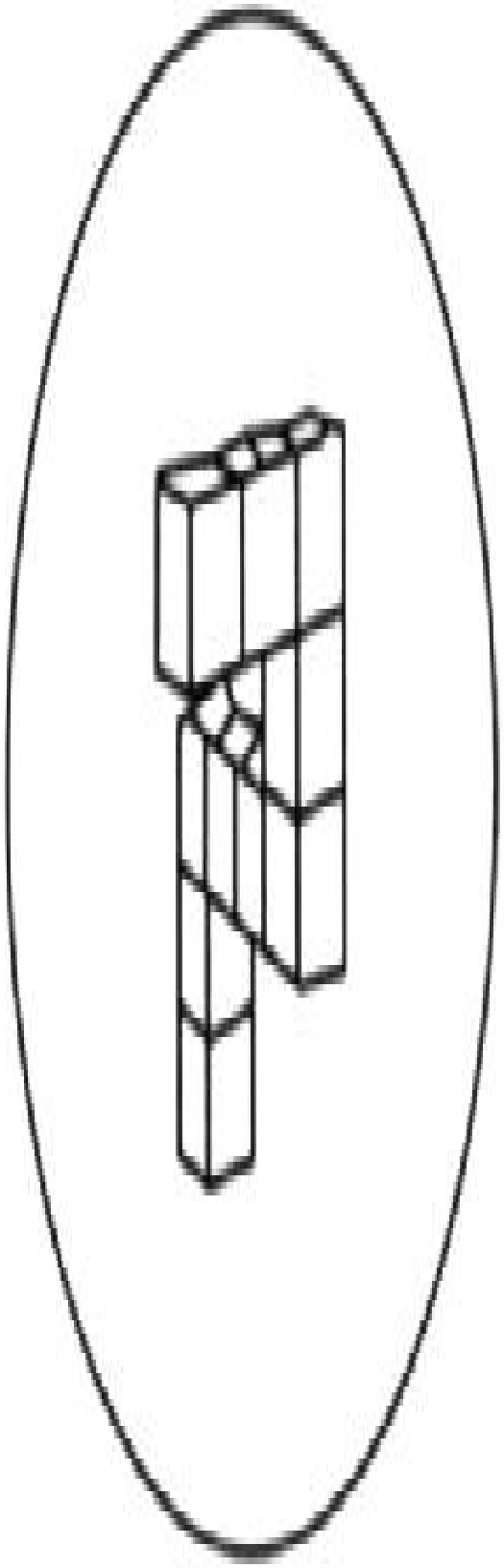
participantes pueden interpretar la tarea de formas bastante diferentes de la que el experimentador tiene en mente. El experimentador puede utilizar materiales o procedimientos inadecuados, o no conseguir controlar el experimento de forma adecuada. Las instrucciones explican la tarea, pero los participantes piensan por sí mismos. Este punto me resulta muy evidente cuando leo a un participante las instrucciones de un experimento, y dice: “Sí, pero, ¿qué es lo que quiere que haga realmente?” Así que, el primer riesgo potencial es que un experimento no evalúe aquello para lo que el experimentador lo diseñó.

Un segundo riesgo se puede ilustrar con un ejemplo de los días en que los ingleses preparaban el té poniendo hojas sueltas en una tetera. ¿Se debería poner la leche en la taza antes o después de haber echado en ella el té? La controversia no tenía fin. No parecía tener más importancia que el hecho de si se debería cocer un huevo apoyado en el extremo ancho o en el estrecho. Pero, Sir Ronald Fisher, el matemático y biólogo, encontró una mujer que afirmaba que a ella le importaba porque podía distinguir si la leche se había echado en primer o en segundo lugar. Fisher era escéptico, así que realizó un experimento. Preparó diez tazas de té, en cinco de ellas puso primero la leche y en las otras cinco puso primero el té. Sin saber cuál era cada una, la mujer tomó un sorbo de cada una en un orden aleatorio y dio su opinión. Acertó nueve veces sobre diez. ¿Fue capaz de notar la diferencia o tuvo suerte? Este dilema se les presenta a todos los experimentadores que obtienen resultados que parecen corroborar sus predicciones. ¿La diferencia es real o tienen suerte? Para valorar esta posibilidad, los psicólogos llevan a cabo una prueba estadística de “significación”. Suponen que no hay diferencia real entre las condiciones experimentales, y calculan la probabilidad de que la diferencia observada se produzca por casualidad. Por ejemplo, ¿con qué frecuencia podría una persona, por casualidad, acertar nueve veces sobre diez al elegir entre dos alternativas? Fisher dedujo la respuesta a partir del cálculo de probabilidades: más o menos una vez por cada cien experimentos. Eso le llevó a inducir que la mujer podía notar la diferencia, dado lo difícil que era que se produjeran sus resultados por causalidad si lo hubiera hecho por adivinación. Los psicólogos se conforman con una diferencia menos llamativa. Infieren que un resultado es significativo si se produce por causalidad menos de una vez por cada veinte experimentos, es decir, tiene una probabilidad de producirse por casualidad menor del 5 por ciento. Este nivel de significación es arbitrario, pero habitual. Es una inducción, y por tanto nada garantiza que la conclusión sea cierta. Ciertamente, como media, por cada veinte experimentos con este nivel de significación, uno da resultados que son consecuencia de la casualidad. Fisher dio algunos sensatos consejos a los experimentadores, que

ellos rechazan bajo su responsabilidad. Una prueba de significación solo les dice qué ignorar –experimentos que no consiguen llegar a resultados significativos–. Los expertos han considerado un resultado como real cuando lo pueden replicar. En estadística existen diferentes escuelas de pensamiento, pero este es el criterio que yo sigo. Los resultados experimentales de los que voy a informar son significativos y replicables.

* * *

Las imágenes se derivan de modelos. Las pruebas para esta afirmación surgen de un famoso estudio ideado por el psicólogo Roger Shepard. La idea se le ocurrió en un sueño. Los participantes veían dos dibujos de una figura “sin sentido” construida con bloques pegados formando un objeto rígido con uniones en ángulo recto. En la Figura 2.4 se muestra un ejemplo. Su tarea era decidir si los dibujos representaban un mismo y único objeto. Llevaban a cabo la tarea mediante una rotación mental del objeto del primer dibujo de manera que pudieran superponerlo con el objeto del segundo dibujo. Cuanto mayor fuera el ángulo que tenían que rotar el objeto, más tiempo suponía la tarea. Hacer una rotación mental de 60º les supuso una media de más o menos un segundo. Aproximadamente les supuso el mismo tiempo si la rotación era en el plano del dibujo que si era en tres dimensiones. Rotar un objeto en el plano del dibujo es como estar rotando el propio dibujo cuando está encima de la mesa. Pero, como muestra la figura, rotar el objeto en tres dimensiones es como girar el objeto acercándolo o alejándolo de nosotros. La tarea no se puede hacer con una imagen bidimensional, sino solo con un modelo del objeto interno y tridimensional. Nuestro sistema visual construye un modelo mental del objeto del primer dibujo, localiza su eje imaginario principal y rota el objeto en torno a ese eje para intentar llevar el modelo a una alineación con un modelo del objeto del segundo dibujo.



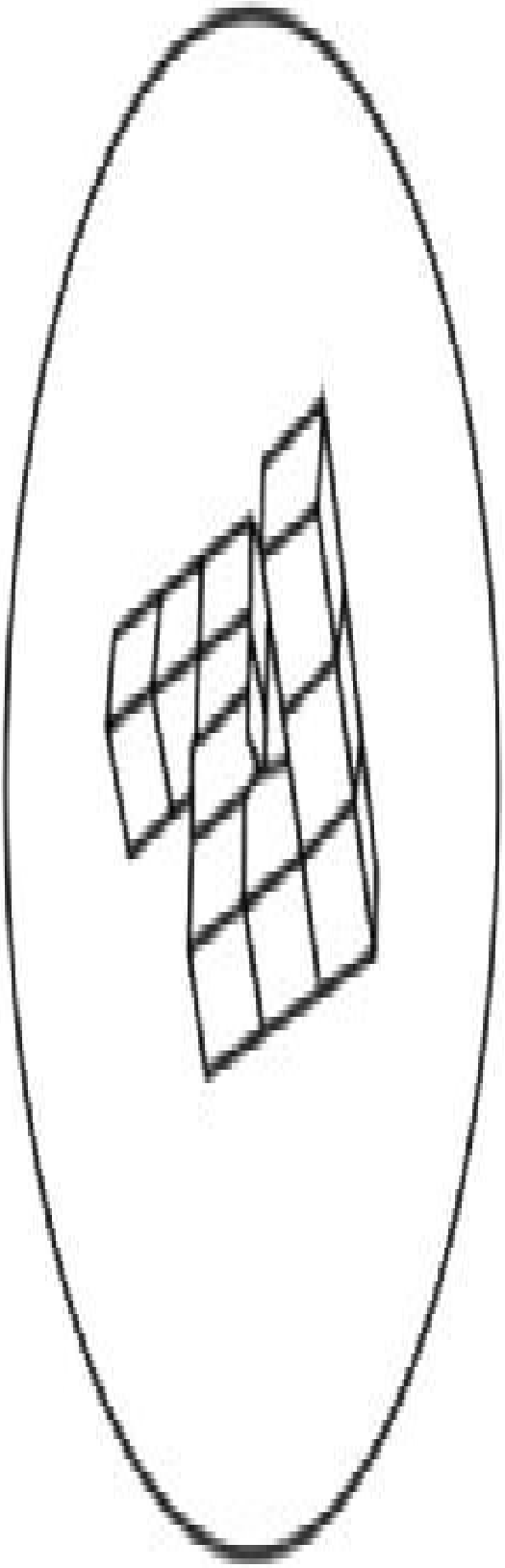


Fig. 2.4. ¿Representan el mismo objeto estas dos figuras (según Shepard)?

Consideremos el sencillo sistema mecánico que se muestra en la Figura 2.5. Si la manivela de la izquierda se hace girar en la dirección mostrada por la flecha, ¿en qué dirección se mueve el eje de la derecha, A o B? La psicóloga Mary Hegarty ha mostrado que cuando tenemos que resolver problemas de este tipo, ponemos en movimiento primero una parte (la manivela), deducimos las consecuencias respecto al siguiente elemento (la polea que está en la misma varilla que la manivela), deducimos las consecuencias respecto al siguiente elemento y así sucesivamente, hasta que o bien perdemos la pista a la rotación o bien llegamos a la solución –el eje de la derecha gira en la dirección A–. El modelo tridimensional subyacente para el sistema se pone en movimiento de una forma cinemática, pero solo elemento por elemento. Un ser sobrehumano pondría en movimiento todo el aparato a la vez. Se decía que el gran ingeniero eléctrico Nikola Tesla tenía esta capacidad y la utilizaba para imaginar qué rodamiento de una de sus dinamos se gastaría antes. Unos tenemos más habilidad que otros para la animación mental. La mayoría de nosotros podemos mover nuestros modelos mentales solo pieza a pieza.

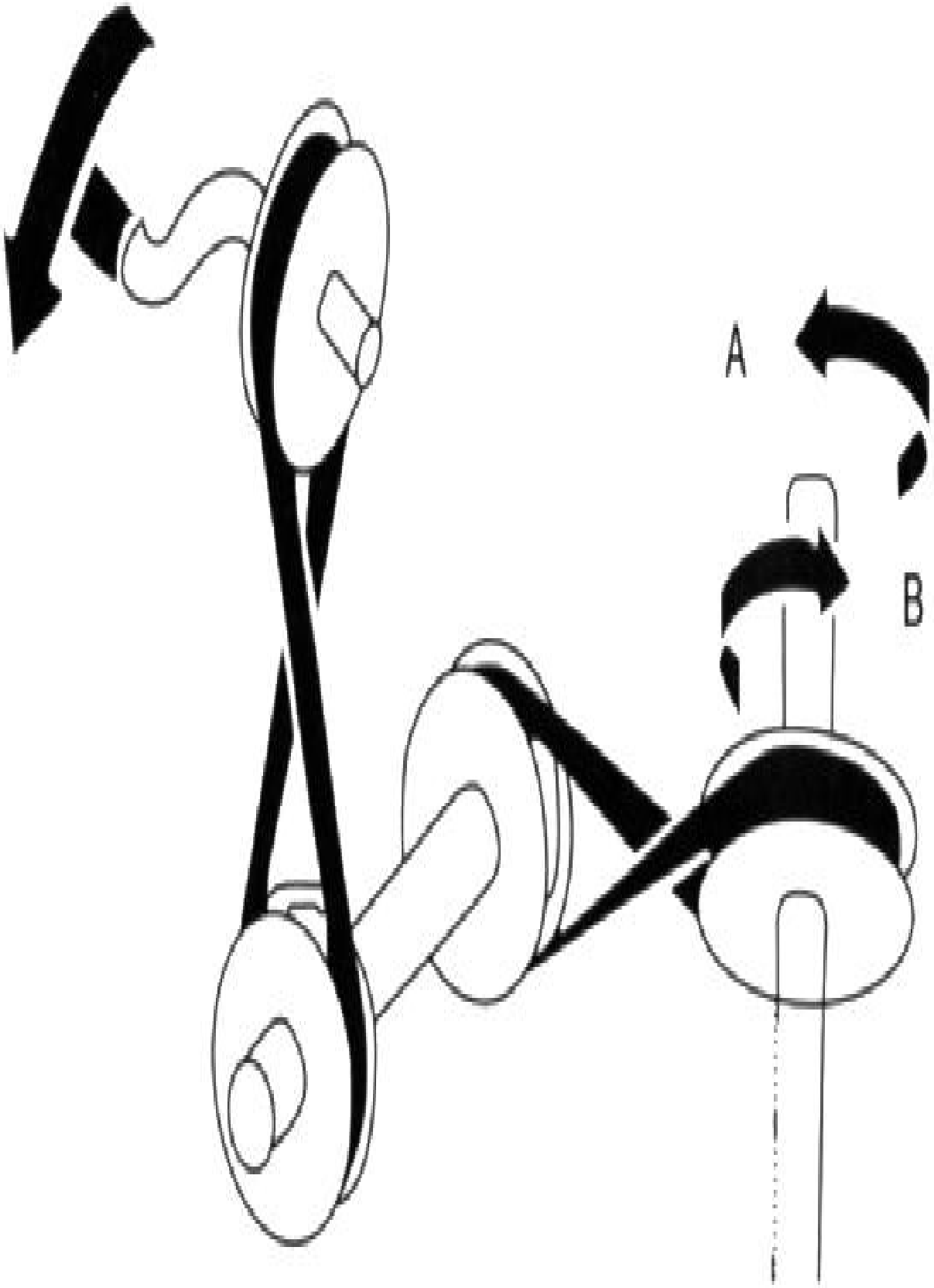

A
B

Fig. 2.5. Si se gira en el sentido de la flecha la manivela de la izquierda, ¿en qué dirección gira el eje de la derecha (según Hegarty)?

En un tercer estudio muy conocido, Stephen Kosslyn pidió a los participantes que revisaran con atención desde un punto a otro una imagen de un mapa sencillo que habían memorizado. El tiempo que les supuso dependía de la distancia que tenían que revisar. Muchos investigadores consideraron que este tipo de estudio implica que la imagen mental es un medio de representación diferente, y Kosslyn sugería que todo conocimiento de alto nivel, incluido el razonamiento, depende de imágenes. Sostenía que las imágenes son vistas de los objetos, y las simuló como disposiciones bidimensionales en un programa de ordenador. En psicología es habitual una línea similar de especulación: la mente evita las representaciones abstractas en favor de representaciones que tienen sus raíces en la percepción, como por ejemplo las imágenes visuales o cinestésicas.

Algunos escépticos, como el científico cognitivo Zenon Pylyshyn, sostienen que las imágenes son "epifenomenológicas" –una palabra de Thomas Huxley para algo que no desempeña un papel causal en un sistema, como el silbido de una máquina de vapor–. Podemos tener intensas imágenes visuales, pero para Pylyshyn no desempeñan un papel causal en nuestro pensamiento. Lo que sí desempeña un papel causal son las computaciones formales que nuestra mente lleva a cabo mediante un único medio de representación, las llamadas "representaciones proposicionales", que son expresiones en un lenguaje mental. Esta afirmación encaja con la opinión de que el razonamiento depende de reglas formales de inferencia.

Por tanto, ¿son las imágenes fundamentales para el razonamiento o son meros epifenómenos? La solución a la controversia es sutil. Por una parte, todas las representaciones mentales se reducen a impulsos nerviosos y a hechos que se producen en las células cerebrales, exactamente igual que la ejecución de cualquier programa de ordenador se reduce al intercambio de "bits" eléctricos de un registro de memoria a otro. En este sentido, Pylyshyn tiene razón: las imágenes mentales deben poderse reducir a una representación de nivel inferior basada en procesos neuronales que tienen lugar en el cerebro. Por otra parte, la imagen mental es una forma distinta de representación mental, y, como veremos

enseguida, puede desempeñar un papel causal en el pensamiento, aunque quizá no el papel en el que pensaban sus defensores.

* * *

Las imágenes visuales son iconos. Su apariencia corresponde a la estructura del objeto o escenario que representan. Algunos teóricos, como he mencionado, dudan si existe alguna representación mental que no tenga sus raíces en la experiencia sensorial. Esta hipótesis es muy recomendable. Es fácil infravalorar el poder de las representaciones perceptivas. Ciertamente, podemos tener una imagen visual de, pongamos por caso, parte del escenario descrito en las orientaciones de Bernad Shaw. Y, de acuerdo con la teoría de modelos, esta imagen es una proyección de un modelo tridimensional subyacente. Pero, no todo puede representarse con un modelo, sea una imagen o un modelo.

Ninguna imagen sola puede captar el contenido de una negación, como por ejemplo:

El mueble de cajones no está detrás del piano.

Podríamos superponer una imagen con una gran cruz roja en nuestra imagen con el mueble de cajones detrás del piano. Pero tendríamos que saber que la gran cruz roja significa negación y no hay nada en la propia imagen que nos lo diga. Asimismo, tendríamos que saber el significado de negación: si una proposición es verdadera entonces su negación es falsa, y viceversa. De nuevo, no hay nada en la imagen que capture ese significado. Podríamos representar la negación mediante imágenes de todas las posibilidades afirmativas. El mueble de cajones no está detrás del piano, y por tanto formamos imágenes con el mueble delante del piano, a la izquierda del piano, a la derecha del piano, debajo del piano, y así sucesivamente. A veces pensamos en estas diversas posibilidades, y nos damos cuenta de que solo una de ellas encaja con la realidad. Para captar su equivalencia con la negación, necesitamos pensar en una disyunción de todas las

diferentes posibilidades. Pero nada en la secuencia de imágenes nos dice que representen una disyunción de la forma: A o B o C o...

Supongamos que pensamos en dos representaciones diferentes. ¿Qué significa el que aparezcan a la vez? Una interpretación es que las dos forman parte de una conjunción: ambas son verdaderas, lo mismo que una proposición conjuntiva (que está formada con “y”) expresa que sus dos proposiciones son verdaderas. Otra interpretación es que las dos forman parte de una disyunción: una de ellas es verdadera, lo mismo que una proposición disyuntiva (que está formada con “o”) expresa que una de sus proposiciones es verdadera. El gran lógico americano del siglo diecinueve Charles S. Peirce diseñó dos sistemas independientes de diagramas para el razonamiento, un sistema utilizaba la interpretación conjuntiva, mientras el otro sistema utilizaba la interpretación disyuntiva. Como comprendió Peirce, la mera inspección de una representación no puede decirnos en ningún caso lo que esta significa. No podemos ver qué sistema de interpretación estamos examinando. Asimismo, lo que un modelo mental representa depende tanto del modelo como del sistema que lo interpreta. El mismo modelo puede tener distintas interpretaciones dependiendo del sistema de interpretación.

La moraleja es que las representaciones son icónicas debido a la forma en que el sistema interpretativo las utiliza. Una representación en sí misma no transmite su interpretación. Conceptos como negación, conjunción o disyunción van más allá de un sistema icónico. Exigen símbolos que se refieran a sus significados. Son abstractos: no pueden representarse solamente con imágenes visuales. Sin embargo, son esenciales para el razonamiento, y por ello los modelos deben contener símbolos que les representen.

* * *

¿El razonamiento depende de imágenes o de modelos? Una forma de responder a esta cuestión es hacer que varíe la facilidad para formar imágenes de las premisas en un problema de razonamiento. Si nos apoyamos en imágenes, deberíamos razonar mejor a partir de premisas que sean fáciles de visualizar, mientras que si nos apoyamos en modelos, no deberíamos razonar mejor, o tal

vez deberíamos razonar peor, a partir de esas premisas. Los psicólogos han llevado a cabo una serie de experimentos de este tipo, y a primera vista los resultados parecen caóticos. A veces el razonamiento mejora con materiales que son fáciles de visualizar; a veces no se ve afectado; a veces empeora. Lo que parece haber sucedido es que los experimentadores han confundido imágenes visuales con modelos espaciales. Relaciones como encima o debajo son fáciles de visualizar y fáciles de representar en un modelo espacial. Sin embargo, no todo lo que es fácil de visualizar evoca un modelo espacial. Relaciones como más claro y más oscuro evocan imágenes visuales, pero no modelos espaciales. Si reexaminamos las publicaciones bajo esta lente, conseguimos enfocarlas correctamente. Cuando los experimentadores utilizaron en las premisas relaciones espaciales, el razonamiento mejoró. Pero, cuando utilizaron en las premisas relaciones visuales, el razonamiento o no se vio afectado o empeoró.

Un experimento examinó esta hipótesis comparando relaciones espaciales con relaciones visuales. Un equipo independiente de expertos valoró un gran número de relaciones. Estimaron que relaciones espaciales, tales como arriba y abajo, evocaban tanto imágenes visuales como representaciones espaciales. Pero estimaron que relaciones visuales, tales como más limpio y más sucio, evocaban imágenes visuales más que representaciones espaciales. El experimento examinó a un nuevo grupo de participantes, que tuvieron que evaluar inferencias, como por ejemplo:

El perro está más limpio que el gato.

El mono está más sucio que el gato.

¿Se sigue: El perro está más limpio que el mono?

Estas relaciones visuales ralentizaron las inferencias de los participantes en aproximadamente medio segundo si se comparaban con las relaciones espaciales. La diferencia puede parecer pequeña pero un análisis estadístico

mostró que era muy improbable que se produjera por causalidad, y que representaba una proporción sustancial del tiempo total empleado en hacer una inferencia, que era del orden de dos segundos. Experimentos posteriores replicaron el resultado, y un estudio mostró, como era de esperar, que sujetos ciegos de nacimiento no están impedidos para el razonamiento sobre relaciones visuales, aunque tardan más en razonar en general.

Un experimento posterior utilizó la técnica de la resonancia magnética funcional (RMf) para determinar qué regiones del cerebro están activas durante el razonamiento con estos materiales. En este procedimiento, se estudió a los participantes mientras tenían la cabeza situada dentro de un gran imán, y las observaciones daban una medida indirecta de las cantidades relativas de oxígeno consumidas por las diferentes regiones del cerebro. Aquellas regiones que son más activas requieren más oxígeno. Los participantes escuchaban las premisas a través de auriculares sonoros de manera que su imagen mental visual no quedara inhibida al tener que leerlas. El razonamiento activaba muchas regiones cerebrales, lo que no era sorprendente. La Figura 2.6, sin embargo, muestra una región del cerebro que solo se activaba en el razonamiento sobre relaciones visuales. Es una región que se sabe que interviene en la visión. Las relaciones visuales pueden ralentizar el razonamiento porque evocan una representación en este área del cerebro, y esta representación es irrelevante para el razonamiento.

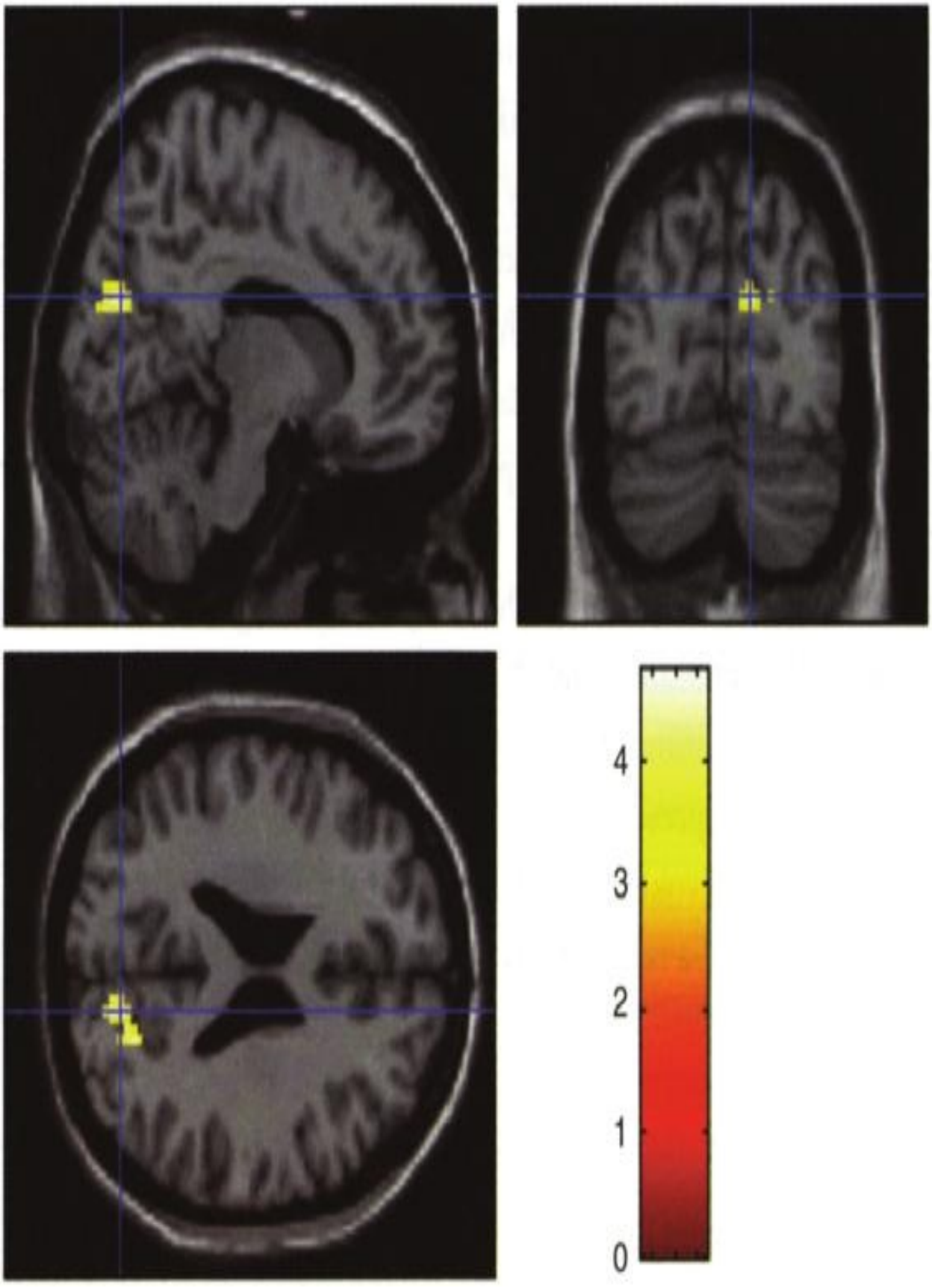

4
3
2
1
0

Fig. 2.6. Tres “lonchas” del cerebro que muestran una región en la parte trasera del cerebro (el córtex visual secundario) que estaba activa al razonar sobre relaciones visuales pero no al hacerlo sobre otras relaciones. El punto de corte de los ejes localiza el área de máxima activación. La escala representa la probabilidad de diferencias casuales, la zona blanca de las figuras corresponde a una probabilidad casual de menos de 1 de cada 20.

* * *

Las imágenes desempeñan un papel causal en el pensamiento porque pueden impedir razonar. Pero, ¿cómo tenemos que interpretar todas las afirmaciones de los científicos sobre el papel de la imagen mental en el pensamiento? En el laboratorio, Ronald Finke y sus colegas han mostrado que la manipulación de imágenes puede ser creativa. Los participantes en su experimento tenían frente a ellos una selección de quince formas: círculo, cuadrado, triángulo, rectángulo, línea horizontal, D, I, L, T, C, J, 8, X, V, P. En cada ensayo, el experimentador nombraba tres formas, y los participantes tenían que cerrar los ojos e imaginar que unían las tres formas para obtener una figura reconocible. Luego dibujaban la figura. Tenían libertad para combinar las partes como quisieran –rotándolas, trasladándolas, superponiéndolas o yuxtaponiéndolas–. Podían cambiar los tamaños de las formas, pero no les estaba permitido deformarlas. Aproximadamente el quince por ciento de los dibujos resultado de su esfuerzo fueron valorados como creativos por un equipo independiente de expertos. La tarea parece depender de un juego visual con las imágenes, y los participantes fueron incapaces de predecir con anterioridad cuál sería la forma resultante.

El juego visual con imágenes de estos estudios recuerda los informes de los científicos sobre su propio uso de las imágenes mentales en la resolución de problemas. La sola visualización, sin embargo, no puede resolver problemas. Si tienen que utilizarse para resolverlos, las imágenes tienen que estar respaldadas por una representación independiente del problema. Como ejemplo, consideremos otra vez el problema de Kekulé. Su objetivo era establecer la estructura molecular del benceno a la luz de la teoría química de valencias –uno

de cuyos fundadores fue él–. La teoría establece las tendencias de los átomos a combinarse con otros átomos. Kekulé sabía que cada molécula de benceno contenía seis átomos de carbono y seis átomos de hidrógeno, y que la valencia del carbono era cuatro, es decir, cada átomo de carbono tiene que combinarse con otros cuatro átomos. La única estructura molecular conocida por entonces eran cadenas de átomos, pero una cadena de seis átomos de carbono requeriría tres átomos de hidrógeno en cada uno de sus extremos y dos átomos de hidrógeno para combinarse con cada uno de los cuatro átomos de la parte central de la cadena. Esto hace un total de catorce átomos de hidrógeno, sin embargo, en la molécula hay solo seis: no son bastantes para cumplir la misión. El problema le dejó sin respuesta hasta que mediante un juego mental con una estructura de cadena formó la imagen de un círculo. Una distribución circular de los átomos sigue sin resolver el problema, porque cada átomo de carbono se une con otros dos átomos de carbono del círculo y con dos átomos de hidrógeno. De nuevo, no había suficientes átomos de hidrógeno para cumplir la misión –solo seis, en lugar de doce–. Kekulé tuvo que hacer la posterior suposición de que uniones alternas entre átomos de carbono tenían una doble valencia, y así cada átomo de carbono del anillo tenía una unión simple con un átomo de carbono y una unión doble con otro. Quedaba sin explicar una valencia en cada átomo de carbono, y era un enlace con un único átomo de hidrógeno. Puesto que los átomos de carbono son idénticos, los enlaces simples y dobles oscilaban de un instante al siguiente. La imagen de los átomos de Kekulé, “todos entretejiéndose y retorciéndose con un movimiento similar al de las serpientes”, puede haber sugerido esta oscilación, pero nadie sabe si la solución se le ocurrió en un sueño.

La manipulación de imágenes es un método para la resolución de problemas visuales. Cambia las apariencias de los objetos. En la resolución de un problema, como la estructura del benceno, el uso de imágenes es parte del proceso. La imagen debe estar relacionada con una representación independiente del problema. Por otra parte, no hay nada que represente el número de átomos y sus valencias excepto lo que aparece en la imagen habitual. La visualización puede producir profundas innovaciones conceptuales pero solo dentro de un sistema que puede representar información más abstracta.

Si puede, forme por favor una imagen visual de esta proposición:

El banco para el pianista se extiende a lo largo de todo el teclado del piano.

Ahora, plantéese estas preguntas:

¿Dónde se encuentra el banco respecto al piano?

¿Cómo están de separados?

¿Qué forma tiene el banco?

Si usted tuviera una imagen vívida, debería ser capaz de responder a estas preguntas. Es decir, es difícil tener una imagen visual del banco y el piano sin representarse sus propiedades y relaciones visuales más destacadas. Por supuesto, mucha gente encuentra difícil formar imágenes, o tenerlas presentes mientras recupera información a partir de ellas. No obstante, una imagen de un escenario pone de manifiesto las características de los objetos de ese escenario desde un determinado punto de vista.

Los modelos mentales son diferentes. Su contenido está disponible para la consciencia, pero no podemos inspeccionarlo como podemos inspeccionar las imágenes. Si rotamos mentalmente un objeto, estamos rotando un modelo tridimensional, pero de lo único que somos conscientes es de una imagen bidimensional que se mueve. Las pruebas respecto a los modelos son por tanto indirectas. A menudo son abstractos y representan una pequeña cantidad de información. Pero su poder reside en su abstracción. Un modelo mental puede representar lo que es común a todas las diferentes formas en que podría producirse una posibilidad. Cada modelo, por tanto, corresponde a un conjunto de posibilidades – definiendo el conjunto mediante lo que contiene el modelo–. Un modelo puede representar que la longitud del banco es igual a la del piano, pero el modelo puede estar libre de los “accidentes” como color, tamaño absoluto, materia, detalle. Puede actualizarse para incluir información sobre estas características. Pero no importa cuántos detalles se añadan a una

descripción de un escenario, la descripción siempre es consistente con una multitud de posibilidades. Un único modelo puede representarlas a todas prescindiendo de los detalles que las distinguen.

* * *

Lo mismo que los pájaros o las abejas, nosotros construimos representaciones del mundo. Estas representaciones son icónicas, es decir, las relaciones y entidades que aparecen en ellas reflejan las que existen en el mundo. Una fuente de representaciones icónicas es la percepción, pero parte del poder del lenguaje es que también él puede llevarnos a construir representaciones icónicas del mundo. Un tipo de icono es una imagen visual. Representa un escenario desde un determinado punto de vista. Otro tipo de icono es un modelo mental, que es una representación tridimensional similar a un modelo real del escenario. También podemos pensar sobre temas que no podemos visualizar, como los conceptos de negación, posibilidad y propiedad. Los modelos son icónicos hasta donde les es posible, pero también contienen símbolos para representar esos conceptos abstractos. Se pueden utilizar para crear imágenes visuales conscientes de escenarios desde un determinado punto de vista. El razonamiento, sin embargo, se basa en modelos antes que en imágenes. Utilizamos las imágenes para pensar sobre cambios en la apariencia de los objetos, pero nuestro juego visual con ellas al resolver problemas depende de una representación independiente del problema. Las imágenes no son por tanto ni fundamentales para el razonamiento ni epifenomenológicas. Son reales, pero no hay que confundirlas con modelos mentales, porque las imágenes se interponen en el camino del razonamiento. William Golding en su maravillosa novela Los herederos describía a una mujer Neanderthal a la que una imagen le dificultaba la tarea de pensar:

“Estoy junto al mar y tengo una imagen. Es una imagen de una imagen. Estoy...” Levantó la cara y frunció el ceño –“pensando”... Se interrumpió, perpleja, la imagen era demasiado vívida y no le encontraba significado.

Notas al pie

[1] Nota de la trad.: El autor utiliza la expresión “dead reckoning”, que significa por estimación (pero también “a ojo de buen cubero”), y considera “dead” como abreviatura de “deduced” [deducido].

CAPÍTULO 3

Modelos de posibilidades: de los trucos de magia a las catástrofes

Hace no mucho perdí un pequeño cuaderno en un hotel de aeropuerto. Acababa de cenar en el restaurante del hotel y regresar a mi habitación, cuando descubrí que el cuaderno no aparecía. Lo primero que debemos hacer cuando hemos perdido algo no es buscar, sino pensar. Tenemos que recordar el último momento en que, con seguridad, tuvimos el objeto perdido y luego repetir nuestros pasos desde ese momento hasta el actual, para localizar todos los sitios en que podría estar. La última vez que había utilizado el cuaderno era en mi habitación del hotel: había hecho una anotación antes de bajar al restaurante. Así que había varias posibilidades: podría estar en algún lugar de la habitación, podría habérseme caído del bolsillo al ir o volver del restaurante o podría estar en el restaurante. Después del pensamiento, la acción. Comprobé la habitación. Llamé al restaurante del hotel: no habían encontrado ningún cuaderno. Pero, solo por asegurarme, bajé y comprobé los alrededores de mi mesa. Comprobé también las escaleras y todo el camino de ida y vuelta a la habitación. Pregunté en el mostrador de recepción. Sin rastro del cuaderno. Era improbable que se me hubiera caído del bolsillo de la chaqueta, y el camarero que me atendió estaba seguro de que no me lo había dejado sobre la mesa. Cuando uno ha eliminado lo imposible, señalaba Sherlock Holmes, cualquier cosa que quede, por improbable que sea, debe ser lo que ha sucedido. Había eliminado todas las posibilidades menos una: a pesar de mi somera búsqueda, el cuaderno debía estar en algún lugar de mi habitación. La siguiente lección en la “ciencia” de encontrar es: pensar de nuevo. ¿Dónde había estado yo mientras escribía las notas en el cuaderno? Había estado tumbado en la cama. Miré debajo de la cama. Miré debajo de las almohadas. Y luego miré entre la cama y la pared. ¡Eureka! Se había caído allí después de que yo hubiera acabado de escribir y me hubiera quedado dormido.

Este ejemplo ilustra un punto importante: cuando pensamos, pensamos sobre lo que es posible. Las principales posibilidades respecto al paradero del cuaderno eran tres: mi habitación, el restaurante o el camino entre ambos. Ser capaz de

pensar sobre las posibilidades es una capacidad envidiable, porque significa que no nos quedamos pegados a lo que tenemos delante de nuestras narices.

Un modelo puede representar una posibilidad real, o una que sabemos que es imaginaria o ficticia. Nuestra mente controla el estatus de los modelos que construye. Tenemos, como dicen los filósofos, “aptitudes proposicionales”. Es decir, sabemos que algunas proposiciones son verdaderas y algunas son falsas; algunas proposiciones las creemos y otras no; y, como punto de partida, estamos dispuestos a considerar algunas proposiciones como hipótesis o asunciones. Nuestros modelos mentales para esas proposiciones tienen etiquetas simbólicas de manera que podamos intentar controlar su estatus. De vez en cuando, en el diagrama de un modelo utilizaré una etiqueta para indicar su estatus. Debería recordar, no obstante, que todos los modelos mentales tienen que tener esas etiquetas. No nos limitamos a pensar sobre posibilidades: representamos el hecho de que son posibilidades.

* * *

Si una persona lanza una moneda, ¿cuántos resultados posibles hay? Es probable que pensemos en dos: cara o cruz. Quizá usted pensó en una tercera posibilidad en la que la moneda se mantiene en equilibrio sobre el canto. Lo normal son dos posibilidades, sin embargo, siempre que la persona que lanza la moneda no sea el famoso estadístico y mago Persi Diaconnis, que podía hacer que la moneda cayese de la forma que él quisiera. Pero, hay otra manera de pensar sobre el lanzamiento de una moneda. Pueden producirse un número infinito de posibilidades. La moneda podría ascender en el aire hasta una altura pequeña, digamos unos quince cm, o hasta una altura considerable, digamos metro y medio; podría subir verticalmente o con un cierto ángulo; podría dar dos vueltas en el aire o doce, podría caer plana en el suelo o girar como una peonza antes de posarse sobre una cara; etc. No obstante, cuando lanzamos una moneda, estamos interesados en dos conjuntos de posibilidades: las que tienen en común que la moneda caiga con la cara hacia arriba y las que tienen en común que la moneda caiga con la cruz hacia arriba. Por consiguiente, cuando nos referimos a una posibilidad nos referimos casi siempre a un conjunto de posibilidades que tienen en común una determinada situación que nos resulta relevante. Es tedioso tener

que seguir escribiendo “un conjunto de posibilidades que tienen en común una determinada situación”, y no lo volveré a hacer. Pero debe usted recordar que cuando me refiero a una posibilidad se trata en realidad de un conjunto de posibilidades.

Nadie negaría que el pensar puede implicar posibilidades. Una afirmación mucho más controvertida es la de que cuando razonamos, nuestro razonamiento se basa siempre en posibilidades. Una alternativa obvia es que se basa, más que en meras posibilidades, en situaciones necesarias –aquellas que es seguro que se producen– o que unas veces se basa en situaciones posibles y otras en situaciones necesarias. Una alternativa más, que discutí en el Capítulo 1, es la de que se basa en formas lógicas que están vacías de contenido. Pero, si el razonamiento se basa en modelos mentales, las posibilidades pasan al primer plano, porque los modelos representan posibilidades. Un corolario es que debería ser sencillo razonar cuando investigamos las consecuencias de una única posibilidad, pero debería ser más difícil cuando tenemos que investigar las consecuencias de más de una posibilidad. Y cuando el razonamiento sea más difícil, nos llevará más tiempo y tendremos más tendencia a equivocarnos. Este es, en pocas palabras, el tema del presente capítulo. Intento mostrar que el razonamiento está basado en posibilidades.

* * *

Supongamos que alguien nos habla del paradero de dos famosos jugadores de cricket que son los capitanes de sus selecciones nacionales:

Sachin está en India o Inzamam está en Pakistán, o ambas cosas.

¿Cuáles son las posibilidades? Cuando se nos pide que las enumeremos, tendemos a escribir tres posibilidades en este orden: Sachin está en India, Inzamam está en Pakistán, Sachin está en India e Inzamam en Pakistán. La disyunción se conoce como “incluyente” porque incluye la posibilidad de que

ambas proposiciones sean ciertas (la tercera posibilidad). Una disyunción "excluyente" excluye este caso, y a menudo se expresa de forma estricta como sigue:

Sachin está en India o bien Inzamam está en Pakistán, pero no ambas cosas.

Descarta la tercera posibilidad, y por tanto enumeramos solo las otras dos posibilidades.

En la vida real, no nos preocupamos mucho por las disyunciones. Decimos, por ejemplo, "el cuaderno está en la habitación o está en el restaurante", y el conocimiento general nos recuerda que el cuaderno no puede estar en los dos sitios. Los lingüistas han argumentado que el significado básico de "o" permite que ambas posibilidades puedan producirse simultáneamente. Este significado puede hacerse explícito con la incorporación de, "o ambas cosas", o por el desagradable uso de "y/o". En cambio, el uso de "o bien", "o si no", o la coletilla "pero no ambas cosas", deja claro que las dos posibilidades no pueden producirse. La diferencia entre los dos tipos de disyunciones se ilustra en una historia sobre el lógico Abraham Fraenkel, que nació en Alemania pero se fue a vivir a Israel. Allí cogió un autobús que tenía su hora de salida de la estación a las 9 a.m. A las nueve y cinco todavía no había salido, y Fraenkel le preguntó al conductor por qué no había salido el autobús, agitando el horario ante él. "¿Qué es usted, preguntó el conductor, alemán o profesor?" Él replicó: "Su 'o' es ¿incluyente o excluyente? Si es incluyente, mi respuesta es sí; pero si es excluyente, es no, soy las dos cosas. "

Si usted entendió mi descripción sobre la pérdida del cuaderno, pensó en tres lugares posibles en relación con su paradero. No tuvo que pensar en todos al mismo tiempo, sino uno por uno. No es fácil tener presente más de una posibilidad a la vez. Es inquietante e impone una carga al sistema de memoria que se conoce como "memoria de trabajo" y es donde mantenemos las cosas mientras pensamos sobre ellas. La memoria de trabajo nunca retiene muchas cosas al mismo tiempo. Es un cuello de botella en la inteligencia. Ciertamente, el punto fundamental respecto a la potencia computacional es la capacidad para retener en la memoria los resultados de las computaciones intermedias. Dado un

pequeño número de operaciones básicas, la potencia de un sistema computacional para hacer computaciones complicadas depende solamente de la capacidad de su memoria de trabajo. Por ejemplo, no podemos hacer multiplicaciones largas sin la memoria de trabajo o un sustituto externo, como papel y lápiz.

Para experimentar por usted mismo la situación de “sobrecarga de memoria”, intente resolver este problema:

June está en Gales o Charles está en Escocia, pero no ambas cosas.

Charles está en Escocia o Kate está en Irlanda, pero no ambas cosas.

¿Qué se sigue?, si se sigue algo.

Podría estar usted tentado a dar una respuesta intuitiva rápida basada en la forma superficial de las oraciones:

June está en Gales o Charles está en Escocia o Kate está en Irlanda.

Pero esta conclusión es una inferencia válida solo si “o” permite que se cumplan todas las posibilidades, y desperdicia mucha información, porque es compatible con siete posibilidades. Así que, por favor, piense por un momento en lo que se sigue de las premisas.

Cada premisa es consistente con dos únicas posibilidades. El problema, por supuesto, es combinarlas. Una estrategia es darse cuenta de que una posibilidad es común a ambos conjuntos de posibilidades:

Charles está en Escocia.

Pero, ¿qué ocurre si Charles no está en Escocia? La única posibilidad que está de acuerdo con la primera premisa es:

June está en Gales.

Y la única posibilidad de acuerdo con la segunda premisa es:

Kate está en Irlanda.

Ambas premisas tienen que ser ciertas, y por tanto estas dos posibilidades deben producirse a la vez. En consecuencia, las dos premisas juntas dan dos posibilidades, que he anotado en líneas separadas

Charles está en Escocia

June está en Gales Kate está en Irlanda

Corresponden a la conclusión:

O Charles está en Escocia o bien June está en Gales y Kate está en Irlanda.

En un experimento realizado con una muestra de población general utilizando este problema, el veintiuno por ciento extrajo esta conclusión válida o una equivalente, pero el resto se equivocó.

Los participantes en el experimento trataron también de resolver problemas basados en disyunciones incluyentes como:

June está en Gales o Charles está en Escocia, o ambas cosas.

Charles está en Escocia o Kate está en Irlanda, o ambas cosas.

Ahora hay cinco posibilidades compatibles con las dos premisas, y las he explicado con detalle en las notas al capítulo. Apoyan la conclusión válida:

Charles está en Escocia o June está en Gales y Kate está en Irlanda, o las tres cosas.

Esta inferencia debería ser más difícil porque depende de cinco posibilidades en lugar de las dos que se necesitaban para la inferencia basada en disyunciones excluyentes. Ciertamente, solo el seis por ciento de los participantes extrajo una conclusión válida. Si hubieran utilizado reglas formales para hacer estas inferencias, la segunda inferencia debería haber sido más fácil que la primera, en contra de lo que indican estos resultados. La dificultad prevista varía porque las teorías psicológicas basadas en reglas formales no tienen reglas para enfrentarse de una sola vez a las disyunciones excluyentes, y por eso necesitan dar pasos adicionales. Podríamos añadir una regla formal respecto a ellas, pero sigue sin existir ninguna razón para suponer que debiera ser más fácil de utilizar que la regla respecto a las disyunciones incluyentes.

La mayoría de las personas se equivoca con ambas inferencias, y podría usted preguntarse qué conclusiones extrajeron. Si están tratando de construir modelos de posibilidades hay dos predicciones obvias. La primera es que si se dan cuenta de que hay más de una posibilidad pero son incapaces de discernir lo que se cumple en todas ellas, entonces deberían responder que no existe ninguna conclusión válida. Aproximadamente un tercio de las respuestas eran de este tipo. La segunda predicción es que si pasan por alto una o más posibilidades, entonces su conclusión debería corresponder solo a algunas de las posibilidades compatibles con las premisas. Cuando los participantes llegaron a extraer una conclusión, la mayoría de sus conclusiones eran de este tipo. Ciertamente, los errores más frecuentes eran conclusiones basadas en una única posibilidad compatible con las premisas. Estos errores no se pueden atribuir a conjeturas a ciegas, debido a la escasa probabilidad de que se hagan tantas conjeturas que den lugar a conclusiones compatibles con las premisas. Explicar la naturaleza de las conclusiones erróneas en el caso de que quienes razonan se basen en reglas formales es tan difícil que nadie ha ideado una explicación de ese tipo.

Una forma sencilla de evitar que los sujetos se agobien con las posibilidades es darles una premisa extra, que establece el paradero claro de uno de los sujetos, p. ej.:

June está en Inglaterra.

June está en Gales o Charles está en Escocia, pero no ambas cosas.

Charles está en Escocia o Kate está en Irlanda, pero no ambas cosas.

La interpretación de las dos primeras premisas da lugar a una única posibilidad:

June está en Inglaterra Charles está en Escocia

La combinación de esta posibilidad con las de la tercera premisa da lugar a:

June está en Inglaterra Charles está en Escocia Kate no está en Irlanda

Por esta vía, el número de posibilidades que hay que retener en cada momento se reduce a una. El experimento incluía algunos problemas de este tipo, y resultaron fáciles.

Lance Rips informó de un extraño resultado experimental que no conseguía corroborar la teoría de modelos. Comparaba dos tipos de inferencia, cada uno de los cuales estaba basado en tres premisas, que incluían dos condicionales, es decir, proposiciones basadas en “si”. Uno de los tipos contenía una premisa inicial que era una conjunción, como por ejemplo:

Anna está en Granada y Pablo está en Barcelona.

Si Anna está en Granada entonces Teresa está en Toledo.

Si Pablo está en Barcelona entonces Teresa está en Toledo.

¿Está Teresa en Toledo?

El otro tipo de inferencia contenía una premisa disyuntiva en lugar de la conjunción inicial: Anna está en Granada o Pablo está en Barcelona. La

respuesta correcta para ambas inferencias es, “si”. Pero, dado que una conjunción tiene un solo modelo, y una disyunción tiene múltiples modelos, la inferencia conjuntiva debería ser más fácil que la inferencia disyuntiva. Sin embargo, no había diferencia fiable entre ellas. Juan García Madruga y sus colegas tampoco consiguieron detectar ninguna diferencia al utilizar los materiales anteriores traducidos al castellano. Pero estos investigadores llevaron a cabo algunos estudios posteriores que mostraron lo que había sucedido. Cuando los participantes en un experimento, en lugar de responder a una pregunta sobre una conclusión dada, sacaban sus propias conclusiones a partir de las premisas, los problemas conjuntivos (noventa y cinco por ciento de respuestas correctas) eran más fáciles que los problemas disyuntivos (setenta y siete por ciento de respuestas correctas). Es más difícil generar una conclusión que evaluar una conclusión dada. Asimismo, cuando las premisas conjuntivas y disyuntivas aparecían a continuación de las dos premisas condicionales, las inferencias conjuntivas también eran más fáciles que las inferencias disyuntivas incluso en la evaluación de conclusiones dadas. La tarea es más exigente porque, cuando los participantes llegan a la disyunción final, la memoria de trabajo ya está previamente ocupada por los modelos de los condicionales. Y cuando las premisas se presentaban de una en una, los participantes gastaban más tiempo en leer detenidamente la premisa disyuntiva que en leer la premisa conjuntiva. En este caso no se producía diferencia fiable en la exactitud de las inferencias, pero los participantes tardaban más en evaluar las inferencias disyuntivas que las inferencias conjuntivas. De ahí que, cuando las premisas y la conclusión se presentan juntas, como en el experimento de Rips, la tarea de evaluación parece no ser lo bastante sensible como para detectar la diferencia entre los dos tipos de premisa inicial.

* * *

Supongamos que se hubiera encontrado una cura para todas las enfermedades. ¿Cuál sería una posible consecuencia?

Podría usted tratar de responder a esta pregunta. Una respuesta habitual es que habría menos médicos, o que muchos médicos podrían arruinarse –a menos que se les necesitase para administrar la cura.

Supongamos que hubiera una guerra nuclear. ¿Cuál sería una consecuencia necesaria –algo que es inevitable que suceda–? De nuevo, usted debería tratar de responder a esta pregunta por sí mismo. La respuesta más frecuente es: moriría mucha gente. Las preguntas sobre consecuencias posibles, como mostró un experimento, son más fáciles de responder que las preguntas sobre consecuencias necesarias. Cuando creamos una posible consecuencia, tenemos que pensar en una posible situación en la que se produce la consecuencia, y no importa si hay otras posibles situaciones en las que no se produce. Pero, cuando creamos una consecuencia necesaria, tenemos que pensar en alguna que se produce en todas las posibles situaciones que podemos imaginar. Por tanto, los participantes emplearon algunos segundos más el pensar en consecuencias necesarias, y estaban menos seguros de sus respuestas.

En el caso de premisas que tienen diversos modelos, un único modelo consistente con ellas –un ejemplo– establece una conclusión sobre lo que es posible, pero para establecer una conclusión sobre lo que es necesario deben servir de ejemplo todos los modelos. El caso contrario es el de las refutaciones: un único contraejemplo refuta una conclusión sobre lo que es necesario, mientras que para refutar una conclusión sobre lo que es posible deben servir de contraejemplo todos los modelos. Y por tanto deberíamos deducir más rápidamente una conclusión sobre lo que es posible que una conclusión sobre lo que es necesario, pero deberíamos deducir más rápidamente una conclusión sobre lo que no es necesario que una conclusión sobre lo que no es posible. Esta predicción es crucial para una teoría que considera fundamentales las posibilidades, y ha sido corroborada por varios experimentos. Una advertencia, aprendida a base de cometer errores, es que no sirve de nada preguntar a adultos si algo es posible cuando su necesidad es obvia. En estas circunstancias, rehúsan decir “sí”, y por tanto las inferencias tienen que ser lo bastante difíciles como para que cuando un suceso es posible no sea también obvia su necesidad. He aquí una prueba típica correspondiente a un experimento. Las premisas se refieren a un partido de baloncesto uno contra uno en el que toman parte dos personas:

Si está Allan [en el partido] entonces está Betsy.

Si está Carla entonces no está David.

Los participantes eran más exactos y rápidos al inferir que Betsy podía estar en el partido que al inferir que debía estar en él. Si usted enumera los posibles partidos compatibles con las premisas, descubrirá que hay tres partidos y que Betsy está en todos: Allan contra Betsy, Betsy contra Carla, Betsy contra David.

Una forma de crear un problema para el que la respuesta es “no”, tanto si la pregunta se refiere a lo que es posible como si se refiere a lo que es necesario, es cambiar en las premisas “estar” por “no estar” y viceversa. Las premisas anteriores se transformarían de esa forma en:

Si no está Allan entonces no está Betsy.

Si no está Carla entonces está David.

Ahora los participantes fueron más exactos y más rápidos al inferir que Betsy no necesitaba estar en el partido que al inferir que no podía estar en él. Los únicos partidos posibles son: Allan contra Carla, Allan contra David y Carla contra David.

Los críticos podrían argumentar que estamos sesgados para contestar “sí” a preguntas sobre posibilidad y “no” a preguntas sobre necesidad, y que ese sesgo produce el resultado decisivo. Sin embargo, en el primero de los estudios que he descrito, los participantes crearon más rápidamente las consecuencias posibles que las consecuencias necesarias. Dado que imaginaban el contenido de esas proposiciones, el resultado no se puede explicar por un sesgo que hace responder “sí” a preguntas sobre posibilidades.

* * *

Cuando razonamos pensamos en posibilidades. Esa es la razón por la que resulta tan difícil razonar sobre parejas de proposiciones disyuntivas. Esa es la razón por la que es más fácil razonar a partir de una disyunción excluyente que permite dos posibilidades que razonar a partir de una disyunción incluyente que permite tres posibilidades. Y esa es la razón por la que nuestras conclusiones erróneas tienden a ser compatibles con solo algunas de las posibilidades: pasamos por alto las otras. El problema al razonar es seguir el rastro de las posibilidades, y en capítulos posteriores mostraré que los diagramas que nos ayudan a hacerlo mejoran nuestro razonamiento.

Podría suponer usted que por alguna otra razón desconocida las disyunciones excluyentes son más sencillas que las disyunciones incluyentes. En realidad, no existe esa diferencia. Cuando tratamos de aprender un concepto nuevo las disyunciones excluyentes son más difíciles que las disyunciones incluyentes. También son más difíciles cuando tratamos de crear un diagrama para conectar unos interruptores y una bombilla de manera que la bombilla se encienda con una disyunción de interruptores. Más adelante explicaré cómo. Las disyunciones excluyentes son más fáciles que las incluyentes solamente cuando razonamos a partir de premisas. La diferencia se explica por el número de posibilidades que tenemos que representar.

Las parejas de disyunciones no parecen típicas de la mayoría de los problemas de razonamiento de la vida diaria. Pero los documentos complejos preparados por los abogados llevan con frecuencia a una multiplicación de posibilidades que supera nuestra capacidad para seguirles el rastro. Aquí aparece, por ejemplo, un extracto de un folleto del gobierno del Reino Unido en el que se describe quién podría optar al llamado “subsidio de defunción”, que ayudaba a sufragar el entierro del cónyuge:

El subsidio de defunción se paga si la persona en virtud de cuyas contribuciones [al Servicio Nacional de salud] se reclama el subsidio satisface cualquiera de las condiciones siguientes:

El contribuyente debe haber pagado o tener reconocidas al menos 25 contribuciones de cualquier tipo en cualquier momento entre el 5 de julio de

1948 o la fecha de ingreso en el sistema de prevención, si ha sido posterior, y el 5 de abril de 1985, o la fecha en que alcanzó los 65 años (60 para una mujer), o murió antes de esa edad, sea lo que sea lo que suceda antes; o

Desde el 6 de abril de 1985 el contribuyente debe haber pagado realmente las contribuciones en cualquier año fiscal anterior al año a considerar, con unas ganancias de al menos 25 veces el salario mínimo para ese año. Habitualmente el año a considerar es el año fiscal (referido a la declaración de la renta) en que se produjo la muerte, pero si inmediatamente antes de la muerte la propia persona en virtud de cuyas contribuciones se reclama el subsidio ha muerto o tiene más de 65 años (60 para una mujer), es o el año en que alcanzó esa edad, o el año en que murió, sea lo que sea lo que suceda antes.

Hay tantas posibilidades que empezamos a perder el rastro hasta de quién ha muerto. No obstante, las parejas de premisas disyuntivas son atípicas, y por eso, en los capítulos siguientes, presentaré más pruebas de que razonamos sobre posibilidades. Gran parte de estas pruebas proceden de estudios sobre inferencias más próximas a las que se realizan en la vida diaria.

* * *

En cualquier situación en que existen posibilidades alternativas, la cognición se hace más difícil, sea al razonar, al planificar, al resolver problemas o al tomar decisiones. La tendencia a pasar por alto posibilidades tiene muchas repercusiones, tanto benignas como malignas. Los magos profesionales anticipan cómo vamos a razonar y se apoyan en nuestra incapacidad para pensar en las posibilidades. David Devant, uno de los mejores magos del siglo veinte, inventó una sorprendente ilusión llamada Mascot Moth que desafía nuestra capacidad para pensar en una posibilidad. Una actriz disfrazada como una hermosa polilla llega al escenario. El mago va a abrazarla, pero ella extiende sus alas sobre el rostro y en el momento en que él la rodea con los brazos se desvanece y desaparece instantáneamente, como una vela que se apaga. La idea para la

ilusión, escribió Devant, le llegó en un sueño. Puede que a usted le apetezca pensar sobre cómo es posible crear este efecto. Una pista: la ilusión no se basa en el uso de espejos. La solución la encontrará al final del capítulo.

La incapacidad para pensar en las posibilidades se produce en muchas catástrofes de la vida diaria. Voy a darle solo un ejemplo, pero hay muchos otros. Los operadores de la central nuclear de Three Mile Island infirieron el 28 de marzo de 1979 que la alta temperatura de una válvula de alivio se debía a una fuga. Durante más de una hora pasaron por alto la posibilidad de que la válvula se había bloqueado abierta. Hicieron bien en pensar en esa posibilidad porque poco después el problema provocó que sonaran más de cien alarmas y las agujas de muchos instrumentos se salieran de sus escalas.

La teoría de modelos, a diferencia de otras descripciones del razonamiento, considera como fundamental la representación de posibilidades. Cada modelo mental representa una posibilidad con una forma tan icónica como sea posible, pero es difícil mantener modelos múltiples en la memoria de trabajo. Más modelos implican más trabajo, lo que conduce a tiempos más largos de razonamiento y a una mayor tendencia al error. He presentado algunas pruebas que apoyan esta predicción, y más tarde volveré a tratar el tema, pero ahora necesito considerar los procesos de razonamiento inconscientes.

La solución al truco Mascot moth es que la actriz que representaba el papel de polilla llevaba un vestido que era igual si ella estaba dentro que si no lo estaba. En el escenario, tras ella, se levantaba un tubo negro de unas cuatro pulgadas de diámetro, el vestido estaba conectado por la parte posterior a un cable que se deslizaba dentro del tubo. El cuello del vestido tenía un muelle de acero para servir de percha y el resto de la prenda colgaba de él. El vestido llevaba también atado un carrete de cordón resistente, que Devant tiraba dentro del tubo en uno de sus movimientos hacia la polilla. La actriz se situaba en un lugar preciso del escenario sobre una trampilla. Plegaba sus alas sobre el rostro y las mantenía unidas. Con el pie, golpeaba sobre la puerta de la trampilla, que se deslizaba bajo el escenario. En ese momento el vestido era una concha vacía sostenida por el tubo. Devant se situaba delante del tubo, al lado del vestido vacío con su pie izquierdo detrás de él. Iba a abrazar a la polilla, y a una señal suya, un tramoyista debajo del escenario tiraba del codón resistente, los muelles del vestido se plegaban, y el vestido era arrastrado a través del extremo del tubo vertical. Devant ocultaba el tubo a la audiencia hasta que era retirado. Entonces se apartaba para demostrar que el escenario estaba vacío. La ilusión era brillante

porque no parecía depender de ningún tipo de aparato. En un instante la polilla estaba totalmente a la vista, al instante siguiente se había desvanecido.

Parte II

El mundo en la mente inconsciente

El centro que no puedo encontrar

Lo conoce mi mente inconsciente

Auden

CAPÍTULO 4

La arquitectura mental y el inconsciente

Va usted a hacer un rápido juicio intuitivo sobre dos hombres a partir de sus fotografías. Simplemente mire las fotografías de la Figura 4.1 y decida cuál de los dos hombres le parece más competente.

Como la mayoría de la gente, probablemente ha opinado usted que parece más competente el hombre de la izquierda. Podemos hacer este tipo de inferencias intuitivas rápidas. Mi colega Alex Todorov y sus colaboradores recogieron opiniones sobre fotografías de parejas de candidatos que se enfrentaban entre sí en seiscientas elecciones para el Senado y la Cámara de Representantes de los Estados Unidos. Lo que resulta sorprendente es que aquellos a los que los participantes consideraban como más competentes tendían a ganar las elecciones. (Sí, el hombre de la izquierda fue el que resultó elegido.) En un estudio posterior, los participantes examinaron las fotos solamente durante un segundo, y sus valoraciones sobre competencia predijeron dos tercios de los resultados en las elecciones, mientras que sus valoraciones sobre la honestidad y el carisma no lo hicieron. Por supuesto, los sujetos que parecen competentes pueden ser competentes, pero también se dan los contraejemplos. Parece ser que Warren Harding fue elegido presidente de los Estados Unidos solamente por su apariencia física, sin embargo, fue incompetente en el cargo.

Fig. 4.1. ¿Qué persona parece más competente?

A
B

Fig. 4.2. ¿Qué persona parece tener más fuerza como líder?

Ahora, por favor, eche una mirada al siguiente par de imágenes que aparece en la Figura 4.2. ¿Qué hombre parece tener más fuerza como líder? En un experimento sin publicar de Tony Little y sus colegas (en la Universidad de Liverpool), los participantes tendían a elegir al hombre de la izquierda. Consideraban también que sería un mejor líder en tiempos de guerra, mientras que consideraban que el hombre de la derecha era más agradable, más inteligente y un mejor líder en tiempos de paz. Las dos fotos eran de la misma persona. Un programa de ordenador modificó su foto aproximándola a una imagen de Bush (el hombre de la izquierda) o aproximándola a una de Kerry (el hombre de la derecha).

El que las elecciones se puedan decidir en función de la validez de una "cara" es una posibilidad alarmante. Pero lo que a mí me preocupa es la naturaleza de las inferencias intuitivas que nos hacen capaces de valorar la competencia a partir de una instantánea de una cara. No somos conscientes de los mecanismos subyacentes. Son tan profundamente inconscientes que el tumbarnos en el sillón de un psicoanalista y la "libre asociación" no hará que se revelen ante nosotros (ni ante el psicoanalista). Freud propuso al inconsciente como el locus de los procesos de pensamiento irracional que se diferencian de los procesos racionales de razonamiento consciente. Pero las inferencias inconscientes a las que me estoy refiriendo se remontan al erudito alemán del siglo diecinueve Hermann von Helmholtz. Él sostenía que la visión debe depender de inferencias inconscientes que, por ejemplo, causan nuestra percepción de la profundidad. Los mecanismos inconscientes de Helmholtz no pueden penetrar en la consciencia porque son los que construyen su contenido. No podemos captar nuestras introspecciones con sus propios medios.

La idea de dos tipos diferentes de razonamiento se remonta al menos a Pascal, que distinguía entre razonamiento intuitivo y matemático. Muchos psicólogos modernos también han defendido teorías de "proceso dual" de razonamiento. En términos generales, distinguen entre las inferencias intuitivas rápidas y las inferencias deliberadas más lentas. Los datos confirman esta distinción. Lo que está menos claro es si hay otros tipos de razonamiento, y si dependen de

procesos independientes o simplemente en un único proceso que opera con diferentes niveles de eficiencia. Mi intención inmediata es aclarar la distinción entre procesos y representaciones y, posteriormente, en el capítulo siguiente proporcionar una teoría sobre los procesos de inferencia inconsciente. Esta teoría, a su vez, tiene ramificaciones relacionadas con las emociones humanas y con el papel del razonamiento en las enfermedades mentales.

* * *

Realizamos procesos mentales sobre representaciones mentales, y por tanto necesito ser claro respecto a la distinción entre ambos. Consideremos, como ejemplo, cómo razonamos sobre las relaciones espaciales. Se nos da, pongamos por caso, este problema:

La taza está a la derecha del plato.

La cuchara está a la izquierda del plato.

¿Cuál es la relación entre la taza y la cuchara?

Nuestro cerebro contiene una serie jerarquizada de procesadores que pueden enfrentarse a este problema. Al nivel más alto, tomamos una decisión de intentar resolver el problema (o, si fuéramos perezosos, decidiríamos no molestarnos). Si realmente seguimos adelante, leemos las premisas y las entendemos. Esta capacidad depende de procesos que identifican las palabras, recuperan las relaciones gramaticales que existen entre ellas, usan esas relaciones para componer significados y establecen las referencias de las expresiones. Somos totalmente ajenos a cómo funcionan esos procesos, pero, de acuerdo con la teoría

de modelos, utilizamos la significación resultante de las premisas para construir un modelo icónico de la distribución. El proceso de construcción es inconsciente, pero da lugar a una representación, y este modelo mental nos capacita para, mediante otro proceso inconsciente, extraer una conclusión: que la copa está a la derecha de la cuchara. El modelo también nos capacita para hacer de memoria un boceto de la distribución de los objetos. En general, en la mente consciente el mundo es una secuencia de representaciones que son el resultado de un conjunto de procesos, y en la mente inconsciente el mundo es el conjunto de procesos en sí mismos. Contiene también algunas representaciones de las cuales no somos conscientes, como las dos imágenes ligeramente diferentes que forman los dos ojos y que nos permiten ver el mundo en tres dimensiones. En otras palabras, todos los procesos mentales son inconscientes. Solamente somos conscientes, en el mejor de los casos, de algunos de sus resultados. Cuando tenemos una intuición, como la de que un hombre parece más competente que otro, ni siquiera somos conscientes de las premisas –aquellos aspectos de su cara y su expresión que nos hacen llegar a esa conclusión–. Sin embargo, cuando hacemos una deducción consciente, al menos somos conscientes de la significación de sus premisas y de su conclusión.

* * *

Un programa de ordenador escrito por mí simula el uso de modelos en el razonamiento espacial basándose en más de cien procesos independientes. Van desde aquellos que inician la construcción de un modelo hasta los que encuentran la relación particular que se mantiene entre un determinado par de entidades en un modelo espacial. La Figura 4.3 es un diagrama simplificado de la estructura del programa. Como puede usted ver, el programa está organizado de forma jerárquica. Tiene dos módulos principales: uno analiza oraciones y compone sus significados basándose en los de sus partes, y el otro construye modelos y los interpreta. En ambos módulos, sin embargo, los procesos en el nivel jerárquico más alto interaccionan solo con los próximos a ellos, mientras que los de nivel jerárquico más bajo a menudo son utilizados por varios procesos de categoría superior. El resultado, como muestra la figura, es una “forma jerárquica enmarañada” más que una clasificación clara. No diseñé esta forma jerárquica e implementé un programa para darle forma, sino que escribí el

código que analizaría una oración, luego el código para utilizar sus resultados para actualizar los modelos, etc.. La escritura de un programa de ordenador para simular el conocimiento es un proceso dialéctico: el programador busca la parsimonia permitiendo a los procesos de nivel alto recurrir a los procesos de nivel más bajo. Mi programa podría haber acabado teniendo otros muchos diseños, pero todos los que fueran eficientes serían formas jerárquicas enmarañadas.

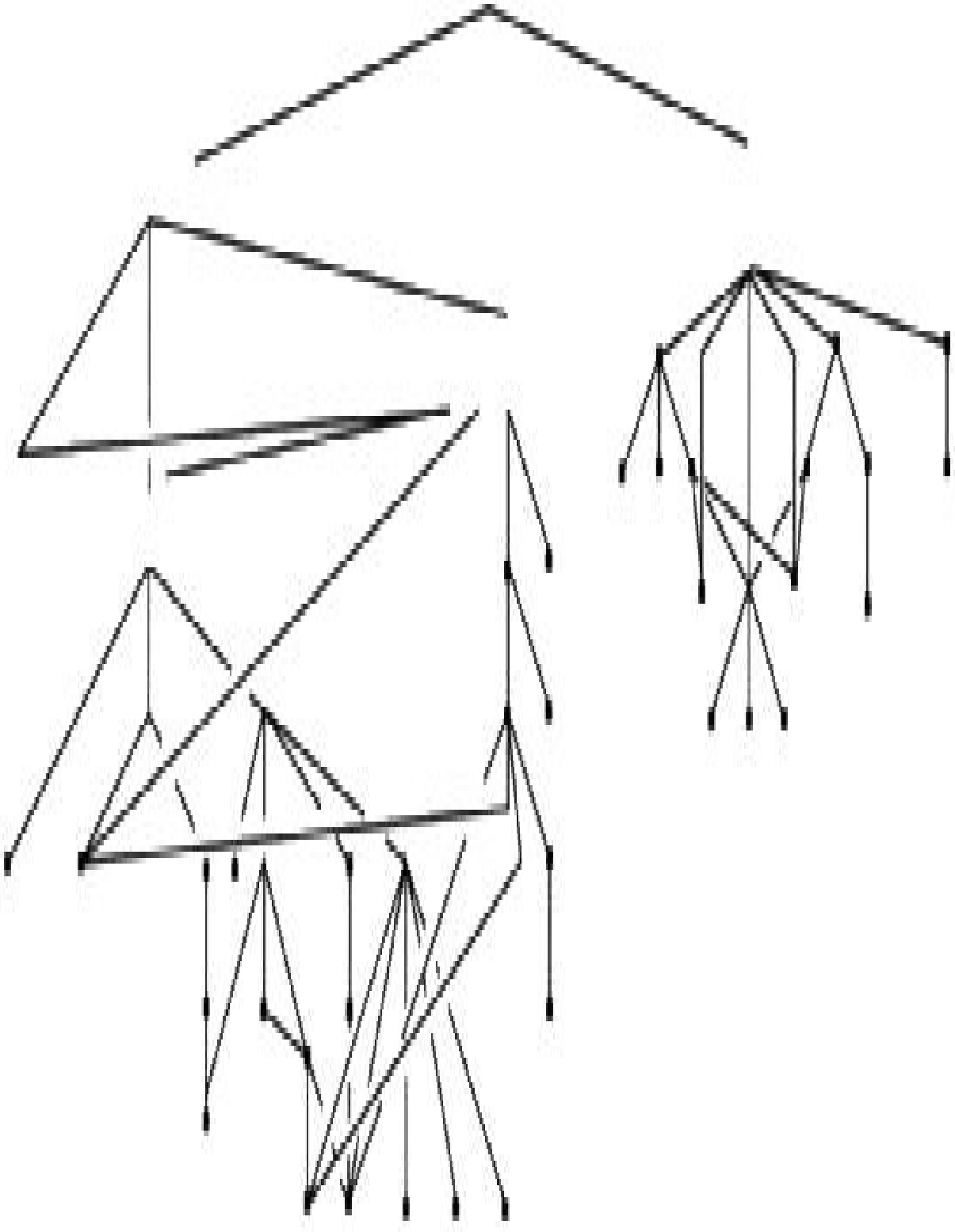


Fig. 4.3. La estructura simplificada de un programa para el razonamiento espacial.

Mi ordenador portátil tiene un único procesador central, que lleva a cabo cada paso del programa partiendo del que está en la cima de la jerarquía e inicia una inferencia y descendiendo hasta los que llevan a cabo operaciones triviales. Por consiguiente, la forma jerárquica de la Figura 4.3 representa cómo un proceso que está ejecutando el procesador central llama a otro, que llama a otro, y así una vez y otra. En cambio, el cerebro contiene miles de millones de diminutos procesadores –células nerviosas– que operan unidas en paralelo. Cada una de estas células nerviosas individuales es lenta al llevar a cabo un proceso, en comparación con el procesador central de mi ordenador. Pero su paralelismo les permite asignar tareas complejas a módulos independientes que trabajan de forma simultánea, de manera que la respuesta global del cerebro es rápida. Una arquitectura de este tipo también es ventajosa para el control de cualquier sistema complicado, como por ejemplo una cadena de montaje automatizada, en la que es necesario llevar a cabo en paralelo una serie de procesos.

El diseño más general para la computación paralela es una red de procesadores conectados a todos los demás de manera que puedan intercambiar instrucciones y datos. Un procesador no puede ver los procesos internos de otro, o influir directamente en ellos. No existe un reloj central que sincronice los procesos. En vez de eso, cada procesador se pone en marcha tan pronto como recibe datos con los que llevar a cabo sus computaciones. Existen muchas variaciones a partir de este diseño. Los canales de comunicación pueden permitir el intercambio de mensajes simbólicos complejos entre los procesadores, o pueden permitir solamente un reducido repertorio de posibles señales. Puede haber un gran despliegue de procesadores que llevan a cabo la misma computación pero con datos diferentes –la llamada computación “en red”–. Sea cual sea el diseño, el procesamiento en paralelo no puede computar nada que, en principio, no pueda ser computado por un único procesador de la serie. Esta es la razón por la que a mi ordenador le es posible simular el comportamiento de un sistema de procesamiento en paralelo.

La computación en paralelo puede acelerar mucho las computaciones. Cada

procesador independiente puede especializarse en determinadas computaciones, y la fiabilidad se puede comprobar en paralelo mediante otros procesadores. El sistema es sólido y resistente a los daños. Sin embargo, las inferencias que producen muchas posibilidades, tales como las del capítulo anterior, son inabarcables. Este término técnico significa que con un número lo suficientemente grande de premisas, cualquier sistema de razonamiento se verá obligado a agotar su tiempo y su memoria antes de poder representar todas las posibilidades. Máquinas tan grandes como el universo trabajando a la velocidad de la luz seguirían tardando miles de millones de años en hacer una inferencia basada en miles de premisas. Y, lamentablemente, el procesamiento en paralelo no convierte esas inferencias en inferencias abarcables. Solamente pospone la crisis. La inabarcabilidad es a la computación lo que la doctrina de Malthus sobre el crecimiento de la población es a la civilización. El índice de reproducción humano crece de manera exponencial, mientras el índice de producción de alimentos crece solo en una progresión aritmética. Resultado: hambrunas. El tiempo requerido para computar una conclusión para una inferencia crece de manera exponencial con el número de premisas, mientras el tiempo disponible para esperar una respuesta no lo hace. Resultado: inabarcabilidad.

El paralelismo tiene ventajas, pero también tiene peligros. Si un procesador está esperando información de un segundo procesador, que a su vez está esperando información del primero, los dos se bloquearán juntos en un “abrazo mortal”. De la misma forma, si un procesador dice avanzar y otro procesador dice retroceder, qué tiene que hacer el pobre organismo –¿partirse por la mitad?–. Un remedio para estos peligros es promocionar a un procesador para que supervise a otros –que arbitre en los conflictos, y esté por encima de los abrazos mortales–. La reproducción de este diseño a gran escala lleva a una jerarquía de procesadores.

Una ventaja de una organización jerárquica es que el controlador global ve la situación “global”. Envía órdenes generales, las cuales se filtran hacia los niveles inferiores de la jerarquía, y van siendo depuradas hasta convertirse en instrucciones precisas. Los neurólogos del siglo diecinueve postulaban este tipo de arquitectura cerebral con el fin de explicar cómo el cerebro gobierna la actuación experta. La decisión de, pongamos por caso, montar en bicicleta pasa a los centros cerebrales que gobiernan los módulos independientes para la destreza, tales como pedalear o conducir, que a su vez dan instrucciones a los centros inferiores que controlan el movimiento de las extremidades, etc., descendiendo en la ruta que llega hasta los nervios que controlan la contracción

de los músculos. Cada centro está computando en paralelo con los otros, y así controlan los hechos que se producen simultáneamente cuando estamos montando en bicicleta, admirando el paisaje y silbando el tema inicial de la Sinfonía Pastoral de Beethoven. Otros centros nos permiten hacer inferencias a partir de proposiciones. Como mi programa de ordenador, reconocen las palabras, analizan las oraciones, componen sus significados y manipulan los modelos. Cada una de estas actividades exige delicadas interacciones entre los procesadores. Deben operar en paralelo, como los trabajadores de una línea de producción de una fábrica, recogiendo que son su materia prima y transformándolas en el producto final que es una inferencia. Somos conscientes de entender el significado de las premisas, quizá de una imagen o dos, y de la conclusión que extraemos. Nuestra consciencia equivale a los resultados de los procesadores de alto nivel, y esas representaciones mentales se mantienen en la memoria de trabajo.

* * *

La consciencia puede deber su origen a la aparición de un controlador de alto nivel en la jerarquía de procesadores. Ese controlador establece metas para los procesadores de nivel inferior, los utiliza para formular planes, recibe información de ellos y monitoriza su actuación. Sus instrucciones se formulan como mensajes, equivalentes, pongamos por caso, a voy a montar en bicicleta, y no como instrucciones detalladas sobre qué músculos contraer. Estas instrucciones se formulan mucho más abajo en la jerarquía. A la inversa, la información que recibe el controlador de nivel alto es compleja, como por ejemplo un modelo del mundo basado en la visión. Nuestra experiencia de la realidad visual es una consecuencia de la arquitectura de la mente. Si pudiéramos realizar una introspección de los procesos que la construyen, lo vívido del contacto con el mundo se vendría abajo. Podríamos dudar de que el proceso fuera verídico, lo que podría ser fatal en caso de peligro. Cuando empezamos a cruzar la calle, queremos ser conscientes de que se nos viene encima una moto, no de los procesos internos de nuestra percepción visual. Supongamos que repetimos esta arquitectura en un robot autónomo. Él también tendría un controlador de nivel alto que ve la situación como un todo, y toma las grandes decisiones. Su controlador sería un equivalente funcional de nuestra

consciencia. ¿Sería consciente el robot? Es difícil decirlo, y lo peor es que no hay una forma obvia de averiguarlo.

Algunos procesos mentales construyen el contenido de la consciencia; otros están en lo profundo del inconsciente. ¿Cuál es la mayor diferencia entre ellos? La hipótesis que yo propongo es que la memoria de trabajo está disponible solamente para aquellos procesos que dan lugar a representaciones conscientes. La memoria de trabajo contiene unas veces una meta, otras un modelo del mundo, otras una señal del medio interno del cuerpo, como por ejemplo hambre o sed, y otras un recuerdo a largo plazo. De estos contenidos es de los que somos conscientes en cualquier momento en que acometemos nuestros asuntos en la vida. Son un resultado de aquello a lo que nos referimos como pensamiento "consciente", aunque depende de procesos que están fuera de nuestra consciencia.

Tenemos conocimiento de nosotros mismos. Sabemos qué tipo de persona somos. Conocemos, hasta un cierto punto, nuestros gustos y preferencias, nuestras capacidades y discapacidades, nuestro pasado y nuestro futuro probable. Tenemos un sentido del "yo", un sentido de nuestra propia integridad, continuidad e individualidad. Tenemos acceso, no a los trabajos internos de nuestra mente, pero sí a sus resultados de alto nivel. Y podemos reflexionar sobre todos estos asuntos cuando tratamos de tomar una decisión. El tomar una decisión intencional depende en parte de nuestra consciencia de que podemos tomar decisiones intencionales. Estas autorreflexiones son nuestra forma de entendernos a nosotros mismos.

¿Qué significaría el decir que un sistema computacional se entiende a sí mismo?

Una respuesta sería que tiene acceso a una descripción completa de sí mismo, y que podría utilizar esta descripción para, pongamos por caso, formular planes respecto a lo que debería hacer. En este caso, la descripción no sería completa a menos que incluyera también este uso respecto a sí misma. En esta idea acechan indicios de que entramos en un círculo vicioso. Nosotros, sin embargo, no tenemos acceso a descripciones completas de nosotros mismos. La naturaleza de la arquitectura mental excluye esa posibilidad. Tenemos acceso, en cambio, a modelos mentales de nosotros mismos. Desarrollamos esos modelos a partir de la infancia, y hay ciertas personas que desempeñaron un papel privilegiado en su desarrollo –nuestros padres, nuestros mentores, nuestros iguales–. Escuchamos lo que tenían que decir, y observamos sus acciones. Observamos también

nuestras propias acciones, pensamos sobre nosotros mismos, y nos evaluamos. Y a partir de toda esta información inferimos un modelo de nuestro yo. Pero es un modelo, lo que significa que es incompleto y con tendencia a ser erróneo. También está inacabado. Día a día, continuamos refinando los modelos que tenemos de nosotros mismos. Puede que no cambien mucho una vez que hemos alcanzado la madurez pero ocasionalmente nos autosorprendemos para bien o para mal, y en consecuencia los revisamos. A veces se desvían tanto de la realidad que necesitamos ayuda para devolverlos a ella.

* * *

La teoría de la arquitectura mental que he descrito implica una grieta entre la conducta consciente e inconsciente. Y existe esa grieta. Por ejemplo, si un fotógrafo nos pide que sonriamos, podemos decidir sonreír. Nuestra decisión es consciente, y nuestra expresión es el resultado de una acción voluntaria. Pero si sucede algo divertido podemos sonreír involuntariamente. Esta sonrisa es el resultado de procesos inconscientes que controlan nuestros músculos faciales. Los dos tipos de sonrisa se diferencian de un modo sutil, y un experto puede decir si nuestra sonrisa es deliberada o espontánea. Si estamos en un cóctel podemos, voluntariamente, trasladar nuestra atención desde alguien que nos habla hasta una conversación próxima. Seguimos asintiendo a quien nos habla pero en realidad no le estamos escuchando a él, sino a otra persona. En cambio, si alguien pronuncia nuestro nombre en una conversación próxima, nuestra atención se trasladará involuntariamente. Un procesador de bajo nivel permanece dormido hasta que la secuencia correcta de sonidos del habla –los que corresponden a nuestro nombre– lo despierta, y traslada nuestra atención hasta la otra persona que está hablando.

En la mayoría de las ocasiones podemos dirigir nuestra conducta. Pero a veces, más de las que nos gustaría, somos incapaces de autocontrolarnos. Podemos intentar ser educados con alguien que no nos gusta –esto es parte del modelo de nuestro yo, pero dice algo muy irritante y le contestamos bruscamente–. Podemos intentar hacer nuestros ejercicios físicos por la mañana, pero cuando nos levantamos no tenemos ganas de hacerlos. Nos confortamos con una racionalización –“los hice ayer, y los haré mañana”– y así no los hacemos. Sin

embargo, a diferencia de lo que ocurre con nuestras acciones, nuestro control sobre nuestros sentimientos emocionales no es mayor del que tenemos sobre nuestras sensaciones corporales, como el hambre o la sed.

Un único procesador en serie, similar al de mi ordenador, no puede estar en conflicto. Pero, un sistema paralelo puede estar en conflicto, y por tanto el controlador que está en la cima de la jerarquía en el cerebro puede estar en conflicto con otros procesadores. Dice, realmente: no te rías, es de mala educación. Pero otros procesadores en un nivel más bajo de la jerarquía dicen eso es divertido, ríete. El conflicto no se resuelve por un acto consciente del controlador, sino por la arquitectura del sistema, es decir, por las interacciones entre sus procesadores. Pueden excitar a otro, reclutando procesadores "que piensen lo mismo"; y pueden inhibir a los procesadores opuestos. Un conjunto de procesadores pueden estar interconectados mediante enlaces tanto excitadores como inhibidores, y todo el sistema puede realizar varias repeticiones de las señales antes de situarse, de una forma o de otra, en una ruta estable. Qué ruta siga depende de diversos factores. La edad tiene importancia: los bebés y los niños no eliminan conductas en beneficio del decoro social. Pero también nosotros nos diferenciamos unos de otros. Y cada uno de nosotros varía de una ocasión a otra. Como señaló el gran poeta francés Paul Valéry, el conocimiento reina pero no gobierna. El controlador no tiene el control total y por consiguiente el camino hacia el infierno está lleno de buenas intenciones.

* * *

El cerebro contiene una jerarquía de procesadores que operan en paralelo. Una ventaja del paralelismo es que permite a procesadores lentos enviar respuestas rápidas al compartir entre ellos el trabajo a realizar. Este diseño no elimina la inabarcabilidad de ciertos problemas, pero permite al cerebro manejar una información mucho más amplia que la que puede soportar nuestra memoria de trabajo. No obstante, la memoria de trabajo permite llevar a cabo computaciones potentes a partir de la limitada información que soporta. Una ventaja de una jerarquía es que el controlador general se enfrenta a la situación "global". Para ello, sin embargo, necesita tener acceso a un modelo del sistema como un todo. Como todos los modelos, es incompleto y con tendencia a ser erróneo. No

obstante, nos capacita para actuar por caminos intencionales –a la luz de nuestro conocimiento de que podemos actuar intencionalmente.

La arquitectura jerárquica implica la existencia de una grieta entre conducta consciente e inconsciente. No podemos ser conscientes de las operaciones de los procesos de bajo nivel. Hay conductas, como el sonreír o el trasladar la atención, que pueden ser conscientes o inconscientes. Algunos aspectos de la vida mental, especialmente las emociones, están fuera del control consciente. Como resultado, pueden producirse conflictos entre conductas voluntarias e involuntarias. Los procesos inconscientes no pueden utilizar la memoria de trabajo, y por tanto tienen capacidades computacionales limitadas para procesar sus grandes cantidades de información. Qué son capaces de hacer exactamente es el tema del capítulo siguiente.

CAPÍTULO 5

Intuiciones y razonamiento inconsciente

Y todas nuestras intuiciones se burlan

de la lógica formal del reloj

Auden

El rastreador de camellos que describí en el Capítulo 1 hizo una serie de inferencias extraordinarias. Mediante el examen de las huellas de los camellos, infirió cuántos habían pasado, la tribu de sus conductores, que tres habían sido robados a otra tribu, el punto de partida del viaje y dónde habían abrevado los camellos. Eran inferencias conscientes a partir de signos de las huellas y llevaban a conclusiones conscientes. Por simplicidad, las llamaré inferencias "conscientes", y solo para asegurarme de que la terminología no resulte confusa aclaro que utilizo "inferencia" y "razonamiento" como acompañantes mutuos: el razonamiento da lugar a inferencias, y las inferencias dependen del razonamiento.

Cuando nosotros hacemos inferencias más corrientes, como por ejemplo:

El cuaderno está en mi habitación o está en el restaurante.

En realidad, no está en el restaurante.

Por tanto, está en mi habitación.

También nosotros hacemos una inferencia consciente que da lugar a una conclusión consciente. Lo que tienen en común las inferencias del rastreador de camellos y las nuestras es que dependen de procesos inconscientes. Todo pensamiento consciente lo hace. El proceso que lleva al rastreador de camellos a inferir la identidad de los miembros de la tribu es inconsciente. Mira los rastros y les identifica. Es consciente de las premisas, que incluyen su conocimiento de los diferentes tipos de rastros que dejan los camellos, pero el cómo las combina para inferir que los conductores eran Awamir está fuera de su consciencia consciente. Tampoco nosotros somos conscientes de cómo hicimos una inferencia consciente respecto al cuaderno. No sabemos cómo construimos una representación mental de las premisas, o cómo la utilizamos para formular una conclusión. De lo que somos conscientes es de nuestro intento de razonar, de la significación de las oraciones, de la conclusión que parece seguirse, de una imagen visual, si es que nos basamos en alguna, y quizá de la dificultad relativa de la tarea. En términos de la arquitectura del cerebro, nuestra consciencia corresponde a los resultados [outputs] de unos pocos procesadores que están en la cima de la jerarquía. Podemos explicar por qué la conclusión se sigue de las premisas. Podemos diseñar un diagrama de los dos lugares en los que puede estar el cuaderno, y explicar por qué, dada la premisa que excluye uno de ellos, se sigue que debe estar en el otro. Incluso la construcción de esta explicación se basa en procesos inconscientes, y algunas veces una explicación puede ser una racionalización más que un reflejo de nuestro razonamiento real. Sin embargo, cuando pensamos en voz alta a la vez que razonamos, aquello de lo que somos conscientes parece ser una descripción auténtica de la secuencia de pensamientos que se construyen mediante los procesos inconscientes.

Un corolario de la teoría de la arquitectura mental es que todas las inferencias dependen de procesos inconscientes. De ahí que, en principio, haya cuatro tipos de inferencia, porque las inferencias pueden empezar conscientemente o inconscientemente, y sus conclusiones pueden ser conscientes o inconscientes. Pero, ¿podría una inferencia iniciarse conscientemente y sin embargo tener solo una conclusión inconsciente? Quizá. El mejor ejemplo es un caso de “fuga” emocional en el que somos conscientes, pongamos por caso, de que estamos enfadados y estamos eliminando los signos externos del enfado, pero no somos conscientes de que durante un breve instante nuestra expresión facial ha

traicionado nuestros verdaderos sentimientos. Los otros tres tipos de inferencia también se producen. La inferencia del rastreador de camellos y nuestra inferencia sobre el cuaderno son ejemplos de inferencias conscientes. Pero, ¿qué pasa con las inferencias que son totalmente inconscientes y las inferencias que se inician inconscientemente pero producen conclusiones conscientes? Presentaré ejemplos de todas ellas como precursor de un análisis del razonamiento inconsciente.

* * *

Algunas inferencias son totalmente inconscientes. Se inician inconscientemente, se llevan a cabo inconscientemente y dan lugar a conclusiones inconscientes. Podría usted preguntarse cómo han detectado su aparición los psicólogos. La respuesta es que estas inferencias afectan a nuestra conducta sin darnos cuenta de ello. Varios estudios han demostrado sus efectos, y voy a describir dos casos sorprendentes.

John Bargh y sus colegas llevaron a cabo un maravilloso experimento de este tipo. Los participantes pasaban una prueba sobre su competencia lingüística en la que tenían que construir oraciones a partir de listas de palabras. Por ejemplo, tenían que usar cuatro de estas palabras para construir una oración:

estar [indicador de futuro] sudor solos ellos[1]

Por consiguiente, podrían responder:

Ellos estarán solos. [They will be lonely.]

Los participantes construyeron de esta forma una serie de treinta oraciones. Y ese fue, para ellos, el fin del experimento. Sin embargo, las listas contenían muchas palabras que sugerían el estereotipo de las personas de edad, p. ej., en la lista anterior: “solos”, y otras como “viejo” y “gris”. Como dicen los psicólogos, estas palabras producen un priming[2] respecto al estereotipo, es decir, hacen que sea más activo, aunque los participantes no sean conscientes de su actividad. Después de terminar el experimento, los participantes recorrieron el pasillo hasta el ascensor. Aquellos que habían recibido el priming respecto a la vejez caminaron más lentamente que aquellos que habían llevado a cabo un test de construcción similar que no suscitaba ese priming. Pero los que caminaron más lento no fueron conscientes ni del priming respecto a la vejez, que no incluía palabras relacionadas con la lentitud, ni de los efectos de ese priming sobre su conducta. Y, como mostró un estudio posterior, el priming no afectó a sus emociones, al menos no de una forma consciente. Podría decir usted que no habían hecho una verdadera inferencia, sino una asociación inconsciente. No se lo discutiré, pero trataré de deducir la naturaleza del proceso mental subyacente dentro de un momento.

Otro brillante ejemplo sobre inferencias totalmente inconscientes se produce en los estudios de estereotipos de Claude Steele. El mero hecho de identificar nuestra propia pertenencia a un grupo puede influir en nuestra actuación en una prueba. Al tener que identificarse como miembros de una minoría, como por ejemplo los afroamericanos, para la que existe un estereotipo de bajo rendimiento en el Graduate Record Examination (GRE)[3], se tiende a puntuar peor en este examen que si no ha habido que identificarse. De la misma forma las mujeres obtienen peores resultados en un test de habilidad matemática cuando se produce el priming respecto a su género. Cuando nos enfrentamos a un reto, la mera amenaza de ajustarnos a un estereotipo negativo empeora nuestra actuación. Es un efecto del que no somos conscientes, aunque es sólido y fiable.

¿Podría una inferencia inconsciente producir una conclusión diferente de aquella de la que somos conscientes? La cuestión surge porque tardamos más cuando tenemos que clasificar algo que está en contra de un estereotipo, p. ej., tardamos un quinto de segundo más en asociar “mujer” con “emprendedora” que en asociar “hombre” con ”emprendedor”. Estas diferencias son maleables, pero para algunos contenidos se desarrollan a una edad temprana. En el caso de algunos sujetos, los resultados están en línea con sus actitudes conscientes; en el caso de otros, son una sorpresa. Muestran, por ejemplo, que quienes somos

blancos tendemos a asociar en mayor grado palabras que denotan cualidades negativas con “negros” que con “blancos”. Y el efecto puede producirse aunque, hasta donde sabemos, no tengamos prejuicios raciales. ¿Tenemos un prejuicio inconsciente? Quizá. Pero “prejuicio” puede ser una palabra demasiado fuerte. Mostramos efectos similares cuando asociamos palabras positivas en un mayor grado a “pájaros” que a “árboles”. El fenómeno parece tener que ver con categorías de evaluación generales más que con nuestras actitudes morales respecto a los individuos.

* * *

Las inferencias se pueden producir en un nivel tan bajo de la jerarquía que las premisas para la conclusión pueden resultarnos inaccesibles. Como sucede en los casos de juicios sobre parejas de fotografías del capítulo anterior, la conclusión ni siquiera parece ser el resultado de inferencias. El razonamiento inconsciente de este tipo da lugar a juicios conscientes, pero no podemos explicar las razones que tenemos para ellos, o bien tenemos que intentar reconstruir las razones, con el peligro de que racionalicemos. Estas inferencias inconscientes subyacen en los presentimientos, las intuiciones, las reacciones instintivas, las conjeturas y los insights. Como veremos, a menudo dan lugar a reacciones emocionales, pero se producen también cuando tenemos que controlar sistemas complejos que contienen más variables de las que somos capaces de manejar de forma consciente, como por ejemplo la simulación de una ciudad en un ordenador. Utilizaré el término intuición para referirme a este tipo de inferencias, en las que somos conscientes de una conclusión o juicio pero no de sus premisas.

En la vida cotidiana se producen muchas inferencias intuitivas. A Peirce, el lógico, le robaron el abrigo y un valioso reloj de su camarote en un barco. Interrogó a los probables sospechosos entre la tripulación e identificó al ladrón, pero no pudo explicar las razones para su identificación. Sin embargo, tenía razón y posteriormente recuperó las propiedades robadas. Por supuesto, no todos nuestros presentimientos son correctos, pero, como señaló Peirce, acertamos más veces de las que justifica la casualidad.

Con lo leído hasta ahora en este capítulo, usted mismo puede haber tenido la

intuición de que el razonamiento inconsciente se refiere a las personas. De hecho, algunos psicólogos han sostenido que los rasgos de personalidad son la bisagra sobre la que giran estas inferencias. Pero no siempre han comprobado si se producen inferencias comparables en otros campos. En realidad, también razonamos inconscientemente en relación con entidades inanimadas. Si somos expertos en algún dominio, pongamos por caso la lucha contra el fuego o el atribuir cuadros a sus autores, formamos impresiones inmediatas sobre los fuegos o los cuadros. Decidimos cómo abordar el fuego o quién pintó el cuadro casi sin tener que pensarlo.

Un experimento clásico que demuestra el desarrollo de las intuiciones se debe a Clark L. Hull. Los participantes veían una serie de ideogramas chinos, y su tarea era aprender para cada uno de ellos un nombre sin sentido. Realizaron la tarea con diversos grupos de ideogramas. Cada grupo era diferente y se presentaba en un orden aleatorio diferente, pero se aplicaban los mismos nombres en un grupo y en el siguiente. Los participantes aprendieron a anticipar el nombre correcto de cada ideograma, y sin embargo fueron incapaces de explicar por qué los nombres eran correctos. En realidad, en todos los ideogramas con el mismo nombre existía una forma subyacente común –un “radical”– pero estaba enmascarado por otros elementos presentes en ellos.

Norman Maier llevó a cabo otro estudio clásico sobre intuición en el que el razonamiento inconsciente llevaba a una conclusión consciente. La tarea de los participantes era atar dos cuerdas que colgaban del techo de una habitación. Solo había un problema. Cuando sujetaban una cuerda y caminaban hacia la otra, eran incapaces de alcanzarla. Es bastante fácil pensar en algunas soluciones, p. ej., atar una cuerda a una silla y luego acercar la otra cuerda hasta ella. Una solución concreta, sin embargo, exigía un insight que se escapaba a la mayoría de los participantes. El truco era poner una cuerda en movimiento, y atraparla mientras se sujetaba la otra. Algunos participantes vieron cómo el experimentador caminaba por la habitación y rozaba accidentalmente (a propósito) una de las cuerdas para hacerla moverse. Estos participantes tendían a llegar al insight crucial. Una vez terminado el experimento, solamente uno fue capaz de recordar la pista que les había dado el experimentador. Los demás mencionaron todo tipo de racionalizaciones respecto a su descubrimiento de la solución, sin mencionar la pista. A partir de la aparición de la pista habían hecho una inferencia inconsciente sin percatarse de ella. Cuando tenemos que pensar en voz alta a la vez que intentamos resolver problemas que dependen de este tipo de insight, es menos probable que tengamos éxito. Es como pensar en cómo atarnos los

cordones de los zapatos a la vez que nos los atamos. Nosotros mismos nos hacemos un nudo.

He aquí una inferencia sencilla que puede usted hacer:

Si puso la escultura en la mesa endeble que hay junto a la puerta entonces se habrá venido abajo.

La puso en la mesa.

¿Qué ocurrió?

Usted infirió, espero, que la mesa se vino abajo. Pero, ¿hizo usted alguna otra inferencia? Puede creer que no, pero está equivocado. Fue usted consciente de sus resultados, pero no de que los había inferido. Usted infirió que "ella" en la primera oración se refería a la mesa, no a la escultura o a la puerta. Infirió que "la" en la segunda inferencia se refería a la escultura, no a la puerta. Estas inferencias inconscientes prepararon el camino a la inferencia consciente de que fue la mesa, no la escultura, la que se vino abajo.

Del mismo modo que hay expertos como el rastreador de camellos, que pueden describir las premisas de sus inferencias, también hay otros expertos que no tienen acceso a las bases de su pericia. He aquí un ejemplo. Los apasionados de la música pueden reconocer al compositor de una pieza musical incluso sin haber escuchado antes esa composición concreta. Eso es Mozart, dicen, o eso es Gershwin. Reconocen las secuencias de notas de una melodía como propias de un determinado músico. Como los participantes en el experimento de los ideogramas de Hull, han aprendido a reconocer el estilo de un músico sin un acceso consciente a los principios que subyacen en su capacidad.

En resumen, el razonamiento inconsciente se produce al menos por tres caminos. Puede ser inconsciente como un todo e influir sobre nuestra conducta fuera de

nuestra consciencia. Puede dar lugar a intuiciones de las que somos conscientes aunque no seamos conscientes de inferirlas. Y respalda incluso la más simple de nuestras inferencias conscientes.

* * *

La teoría de la arquitectura mental restringe la potencia del razonamiento inconsciente. No puede utilizar la memoria de trabajo. La memoria de trabajo, como usted recordará, es donde mantenemos presentes las cosas mientras trabajamos sobre ellas. Es el núcleo de la potencia computacional, puesto que mantiene los resultados de las computaciones intermedias. Así que, ¿qué puede hacer el sistema inferencial inconsciente?

He aquí una forma de plantearlo, la que he tomado prestada de la novela satírica de Michael Frayn, The Tin Men. El protagonista que plantea el novelista, Goldwasser, es un periodista inglés que imagina que los artículos muy manidos podrían muy bien escribirlos los ordenadores. Si tiene que escribir otra nueva historia sobre una celebración de la realeza, el artículo completo será un cliché desde el principio hasta el final. Imagina una forma de poder hacerlo utilizando una serie de fichas. Supongamos que tiene que escribir una historia sobre una boda real. Coge la primera ficha, que dice: “Tradicionalmente.” Selecciona a continuación una tarjeta adecuada para bodas, que dice, “las bodas son ocasiones para el regocijo”. Las alternativas son secuencias de tarjetas que cubren historias sobre coronaciones, compromisos matrimoniales, funerales, mayorías de edad, nacimientos, muertes y ordenación de mujeres. Si hubiera estado escribiendo sobre un funeral, podría haber escogido en lugar de esa, una tarjeta que dijera, “son ocasiones para el duelo”. La siguiente tarjeta para las bodas dice, “La boda de X e Y”. Conduce a una elección entre, “no es una excepción”, y “es un claro ejemplo”. En este punto no hay elección, la siguiente tarjeta dice: “ciertamente”. Sea cual sea la historia siempre se llega a ese punto concreto. Y de esa manera, eligiendo al azar entre tarjetas alternativas, Goldwasser puede hilar el resto de la historia:

Es una ocasión especialmente feliz. Pocas veces puede haber habido una pareja

más popular. X se ha ganado un lugar especial en los sentimientos de la nación, y el pueblo británico ya ha tomado cariño a Y. Es especialmente adecuado que..., etc., etc., etc.

Se dará usted cuenta de que el sistema depende del contenido de una memoria a largo plazo –las frases de las fichas–. Podría organizarse de modo que ni siquiera tuviera que recordar si el tema de la oración actual es un personaje masculino o femenino. Podría haber secuencias independientes de tarjetas, una para hombres, otra para mujeres y otra para parejas. Así, si una oración empezara por “Carlos se ha ganado”, habría una secuencia en la que los pronombres serían “él”; si empezara por “Camila”, habría una secuencia en la que los pronombres serían “ella”, y si empezara por “La feliz pareja”, habría una secuencia en la que los pronombres serían “ellos”. Este sistema evita la necesidad de que la memoria de trabajo siga el rastro del género y el número.

Sin memoria de trabajo, nuestro sistema inconsciente para las inferencias intuitivas tiene una organización semejante. En cada punto en que tiene que elegir, el sistema se encuentra en un determinado estado en el que puede llevar a cabo una operación entre un conjunto de alternativas, cada una de las cuales produce un único resultado [output]. Cada estado del sistema permite una elección similar a la que se realiza entre las frases alternativas en una historia sobre la realeza, pero nuestro sistema inconsciente puede permanecer en el mismo estado para llevar a cabo una de las operaciones del mismo conjunto o cambiar a otro estado. A diferencia del sistema de Goldwasser, incluso puede retroceder a un estado anterior dentro de la secuencia. Llega a pararse –quizá con algún resultado para otro sistema– cuando no hay un estado al que pueda pasar. Se mantiene entonces en una pausa hasta ser lanzado a un nuevo estado por una entrada procedente de cualquier punto en la jerarquía de los procesadores.

El sistema inconsciente puede hacer una transición a uno de los estados pertenecientes a un conjunto de posibles nuevos estados, y puede que la elección no la fije el estado actual, del mismo modo que Goldwasser elegía al azar entre las alternativas. Sin embargo, algunas transiciones pueden ser más probables que otras. Ciertamente, las probabilidades pueden depender de la historia reciente de los estados a través de los que ha pasado el sistema. Los matemáticos se refieren a este tipo de sistemas como “máquinas de estadio finito de Markov con límite k”, lo cual es un trabalenguas. (Andrei Markov, un matemático ruso, estudió

estos sistemas.) Para simplificar, me referiré simplemente a un proceso de transiciones de un estado a otro.

La mayor diferencia entre las fichas de Goldwasser y nuestro sistema inconsciente es que, en nuestro sistema, las transiciones tienden más a estar limitadas por el contenido que por las palabras o las frases. Si el sistema es inducido, pongamos por caso, por la proposición: ellos estarán solos, su contenido puede encajar con otras muchas representaciones de la memoria a largo plazo, y de esta forma provocar un priming del concepto de personas ancianas. Si una transición produce una conclusión, entonces una transición posterior puede producir otra conclusión, y así sucesivamente en una corriente de conclusiones que son simples atribuciones de propiedades o instrucciones para la conducta. El contenido de cada conclusión puede servir de ejemplo a otros conceptos o principios en la memoria a largo plazo.

Si el sistema no tiene acceso a la memoria de trabajo, a las inferencias inconscientes les faltará la potencia computacional del razonamiento consciente. No harán frente a las posibilidades alternativas de una disyunción, a las comparaciones entre posibilidades –tales como futuros sucesos alternativos, o posibilidades secuenciales en un orden temporal, o comparaciones entre los hechos reales y una historia alternativa en la que se produjeron otros sucesos–. No harán frente a proposiciones que afectan solamente a algunos de los elementos de un conjunto, p. ej., “Algunas cosas que asustan son peligrosas”. No harán frente a las relaciones que se establecen entre múltiples entidades. Solo contarán hasta un número reducido. No llevarán a cabo una determinada secuencia de operaciones en orden inverso. No incluirán una secuencia de operaciones dentro de otra secuencia, etc., de forma indefinida. Por tanto, no establecerán dependencias de amplio espectro dentro de dependencias de amplio espectro. El sistema de fichas puede asegurarse de que se usan los pronombres adecuados en la historia sobre Carlos, pero las fichas no pueden crear una historia que incluya dentro de ella otra historia sobre Camila, la que a su vez incluya una sobre el príncipe Harry, etc. Nuestras inferencias inconscientes tendrán las mismas limitaciones, porque ese tipo de “inclusiones” exige una memoria de trabajo. El razonamiento inconsciente se ocupará de las posibilidades de una en una, habitualmente de una que exista aquí y ahora. Conducirá de una posibilidad a otra, y dependerá del acceso a ciertos conceptos y principios generales de la memoria a largo plazo, a los que me referiré dentro de un momento. Solo si las conclusiones intermedias llegan a ser conscientes, nuestro razonamiento superará las limitaciones que he enumerado aquí. Pero

también la memoria de trabajo tiene sus limitaciones y por eso nuestro razonamiento consciente es humano, no sobrehumano.

Cuando reconocemos un estilo musical, parece que nos basamos en transiciones. Escribí un programa de ordenador que aprendió las transiciones de los tonos y las duraciones –un par conjuntamente con el siguiente– a partir de un corpus de improvisaciones musicales. Luego utilizó esas transiciones para generar nuevas improvisaciones. Eran simulaciones identificables del estilo de un músico determinado. La improvisación de melodías no parece necesitar memoria de trabajo, aunque sí se necesita para enfrentarse a la estructura a gran escala de una composición musical. No obstante, la notación musical puede funcionar como un sustituto de la memoria de trabajo.

También parecemos basarnos en transiciones al identificar quién pintó un cuadro determinado. Los entendidos pueden identificar un cuadro como un Picasso o un Monet incluso si no lo han visto nunca antes, y pueden hacerlo sin ser capaces de explicar con palabras las señales que utilizan. La habilidad de las palomas como entendidas en arte corrobora la idea de que este tipo de destreza no depende de la potencia computacional de la memoria de trabajo. Aprendieron a distinguir entre pinturas de Monet y Picasso. Su destreza se extendió también a cuadros que nunca antes habían visto. Y cuando vieron cuadros de pintores nuevos hicieron atribuciones sensatas. Si se les presentaba un Renoir o un Cezanne, consideraban que era un Monet. Sin embargo, si se les presentaba un Matisse o un Braque, consideraban que era un Picasso. Por tanto, mostraron una cierta apreciación de los linajes del impresionismo y el arte de comienzos del siglo veinte.

Otra clave respecto al razonamiento sin memoria de trabajo surge de los estudios con bebés, que tienen memorias de trabajo muy limitadas. Como ha demostrado Gary Marcus en un brillante estudio con niños de siete meses, pueden hacer generalizaciones para todos los ejemplos de una categoría abstracta. Los bebés escucharon una serie de sonidos como “ga na ga” y “li ti li”. Todos los sonidos se adaptaban a la estructura abstracta ABA, en la que cada letra indica una sílaba. Otros bebés participantes en el experimento oyeron sonidos en los que la estructura era siempre, ABB, p. ej., “wo, fe, fe”. Después de escuchar estos sonidos por espacio de unos dos minutos, los bebés oyeron sonidos formados con nuevas sílabas que no se habían utilizado en la sesión de entrenamiento. La mitad de los sonidos eran consistentes con la estructura de las oraciones anteriores, y la mitad no. Por ejemplo, el sonido: “wo, fe, fe” sería inconsistente con la estructura ABA. Los niños prestaban más atención a los sonidos nuevos

que eran inconsistentes con la estructura abstracta que habían adquirido durante la fase de entrenamiento. Por tanto, los bebés generalizaban esta estructura a todas las posibles sílabas, incluidas las que no habían escuchado anteriormente durante el experimento. Este resultado desafía a las teorías del aprendizaje que solo representan relaciones entre sílabas concretas, y que carecen de la potencia para generalizar a todo tipo de sílabas.

Algunas transiciones dan lugar a emociones. Por ejemplo, el sistema inconsciente evalúa que una situación es peligrosa, y entonces hace una transición a una emoción. Ve usted algo peligroso, como por ejemplo un coche girando sin control frente a usted. Y se siente asustado y realiza una acción evasiva. En realidad, el sistema ha hecho esta inferencia:

Esto es peligroso.

Por tanto, esto asusta.

Pero, en lugar de una conclusión, obtiene usted una emoción, y lleva a cabo una acción. Un corolario es que podemos ser conscientes de la causa de la emoción, el coche girando, y del miedo en sí mismo. Pero no somos conscientes de la transición desde la una al otro. En el capítulo siguiente volveré a ocuparme de este punto cuando me ocupe de las emociones.

* * *

Nuestro sistema de inferencias inconscientes depende de recuerdos a largo plazo. Probablemente son principios de conocimiento general que hemos adquirido por inducción. Pero los principios en sí mismos pueden ser inconscientes –como las transiciones que el crítico musical ha adquirido en relación con las melodías de Mozart–. Otros aspectos de nuestro conocimiento general pueden tener la forma

de modelos mentales sobre cómo funciona el mundo. Y como verá usted posteriormente a lo largo del libro, a partir de esos eslabones básicos de nuestro conocimiento general podemos hilvanar una descripción causal de un suceso. Podemos utilizarlos para hacer una serie de inferencias elementales, p. ej.:

Viv insultó a Pat.

Por tanto, Pat se enfadó.

Por tanto, Pat perdió los nervios.

Por tanto, Pat fue grosera.

Nuestras intuiciones basadas en inferencias inconscientes pueden entrar en conflicto con las inferencias conscientes. Por ejemplo, considere este problema:

Todos tenemos prejuicios contra la gente que tiene prejuicios.

Anne tiene prejuicios contra Beth.

Por tanto, ¿se deduce que Chuck tiene prejuicios contra Di?

Nuestra intuición es: no, por supuesto que no, porque no se ha dicho nada sobre

Chuck o Di. Pero nuestra habilidad consciente nos permite seguir este argumento: sabiendo que Anne tiene prejuicios contra Beth, Anne es una persona que tiene prejuicios y por tanto todos tenemos prejuicios contra ella. Pero, en ese caso, Di tiene prejuicios contra ella, y por tanto Di es una persona con prejuicios. Por tanto, todos tenemos prejuicios contra Di y esto implica que Chuck también tiene prejuicios contra ella. La inferencia depende de un bucle de operaciones repetidas para actualizar nuestro modelo de la situación, y ese tipo de bucles dependen de la memoria de trabajo. En los años sesenta, los publicistas experimentaron con la publicidad subliminal, lanzando breves imágenes durante las películas para animar a los espectadores a comprar bebidas gaseosas. No funcionó. Las inferencias que necesitamos hacer para comprar una Coca-cola superan la potencia de nuestro razonamiento inconsciente. Exigen planear una secuencia de acciones interdependientes, y ese tipo de planes dependen de la memoria de trabajo.

* * *

Las transiciones se apoyan en parte en principios generales del conocimiento a largo plazo. Estos principios tienden a ser heurísticos –reglas simples de sentido común que las transiciones pueden utilizar con facilidad–. Su contenido se revela en aquellas inferencias que son rápidas, automáticas e involuntarias. He aquí tres ejemplos diferentes:

Oímos una fuerte explosión. Inferimos que algo la ha causado.

Cogemos a alguien por los hombros y grita. Inferimos que le hemos hecho daño.

Un objeto choca con otro, que empieza a moverse. Inferimos que el primer objeto causó el movimiento del segundo objeto. Hacemos esta inferencia si el orden temporal de los sucesos es adecuado, incluso cuando sabemos que estamos

viendo una imagen de ordenador.

Estas inferencias se basan en un conjunto de principios causales plasmados en el sistema inconsciente:

• Todo suceso tiene una causa.

• Los objetos en movimiento son causas.

• Los humanos y los demás animales pueden poner en marcha causas.

• Las causas están en contacto físico con sus efectos.

• Causas similares tienen efectos similares.

Estos principios pueden parecer razonables, pero cuando los examinamos con detalle, algunos son bastante sospechosos. Entonces, ¿por qué los tenemos? ¿Y por qué tenemos otros principios análogos que nos capacitan para hacer inferencias sobre acciones deliberadas e interacciones sociales? La respuesta puede ser que estamos equipados con ellos, porque con frecuencia tenemos que inducir relaciones causales, y la inducción nos exige ir más allá de la información que se nos da. La inducción de las relaciones causales correctas es inabarcable: el número de posibles hipótesis inductivas crece en una proporción exponencial respecto al número de conceptos que pueden ser relevantes. Pero no sabemos qué conceptos son relevantes, y por tanto es fácil invocar los conceptos erróneos. En consecuencia, las restricciones en la búsqueda de posibles causas nos ayudan, y los heurísticos causales que he mencionado son a menudo útiles.

Una manifestación de las restricciones heurísticas aparece en lo que los antropólogos llaman pensamiento “mágico”. Produce concepciones erróneas de causa y efecto. La magia contagiosa atribuye la causación al contacto físico. Por ejemplo, la cola de los peces se agita, así que si comemos pescado, nos agitaremos. Por tanto, los talladores de madera de la isla de Kitawa en el Pacífico no comen pescado, ya que no quieren que sus manos se agiten cuando

tallan la madera. En la raíz de otros tipos de pensamiento mágico está el que las cosas semejantes se comportan de manera semejante, la llamada magia homeopática. Las personas que viven en los barrios bajos son malas –como mínimo son pobres y esto es malo, y por tanto su conducta es mala–, son inmorales. En consecuencia, Dios castiga su maldad con enfermedades como el cólera. La semejanza como base para la inferencia lleva a una concepción de "acción a distancia" –no importa que este principio sea inconsistente con la causación por contacto–. Quiero hacer daño a mi enemigo, y por eso hago una efigie suya y le clavo alfileres. El resultado es que mi enemigo siente dolor. Cuando Georges Orwell era estudiante, él y un amigo hicieron un modelo de cera de otro chico al que odiaban. Partieron el brazo del muñeco. Algunas semanas más tarde el chico se rompió el brazo, y no mucho después falleció de una enfermedad mortal. A Orwell y a su amigo les invadió el sentimiento de culpabilidad. Un biógrafo reciente sugiere que parte de las razones por las que Orwell escribió bajo un pseudónimo (su verdadero nombre era Eric Blair) era que temía que las personas que conocían su verdadero nombre pudieran también hacer magia contra él.

La magia contagiosa y la homeopática pueden actuar conjuntamente. Los miembros de una tribu africana, los Azande, observaron una conducta sorprendente en el mono de los arbustos rojos. Cuando se despierta, se estira como si estuviera sufriendo un ataque epiléptico, pero luego se recupera. Los Azande razonaron que funcionaría de la misma forma un remedio para la epilepsia. Quienes sufrieran la enfermedad deberían comer las cenizas del cráneo de un mono. La causación por contacto les daría los poderes del mono. Causas similares tienen efectos similares, y por tanto ahora ellos también se recuperarían.

El pensamiento mágico restringe la inducción, y eso ocurre en todas las culturas, incluida la nuestra. El pensamiento mágico de los otros nos resulta evidente, pero el nuestro propio es más difícil de percibir, porque lo hemos utilizado para inducir nuestras creencias. Sin embargo, no tenemos que buscar mucho para encontrar ejemplos de pensamiento mágico en nuestra propia cultura. La tendencia a adjudicar causación al contacto físico puede ser irresistible. Apretamos una tecla en el ordenador y la pantalla se oscurece. Inferimos que la causa fue el apretar la tecla. Quizá lo fue. Pero podemos descubrir más tarde que la batería del ordenador se estaba agotando y como consecuencia el propio ordenador se apagó.

Un ejemplo excelente del pensamiento mágico occidental es el uso de la terapia electrocompulsiva (TEC) para curar la depresión. Los doctores observaron que los epilépticos sufren muy raras veces de esquizofrenia. Razonaron que la esquizofrenia podría aliviarse si inducían ataques epilépticos en los pacientes –un paso que está en paralelo con el pensamiento de los Azande–. Por tanto, indujeron la epilepsia en sus pacientes utilizando la TEC. El tratamiento no funcionó. Pero los médicos notaron que aliviaba la depresión, y así se convirtió en un método estándar de tratamiento para los casos severos. Funcionó incluso cuando, sin que los doctores de un hospital de Londres lo supieran, la máquina de TEC estaba estropeada y funcionaba de manera incorrecta.

* * *

La teoría de la arquitectura mental permite que las inferencias se inicien tanto consciente como inconscientemente. Las inferencias conscientes, por supuesto, no son nunca totalmente conscientes. No podemos hacer introspecciones sobre los procesos que llevan de las premisas a la conclusión: si pudiéramos, entonces esos procesos no serían misteriosos. Las inferencias conscientes, sin embargo, tienen acceso a la memoria de trabajo para mantener los resultados de las computaciones intermedias. En cambio, las conjeturas, los presentimientos y las intuiciones surgen de inferencias inconscientes, que son rudimentarias, porque se hacen sin acceder a la memoria de trabajo. El sistema inconsciente hace transiciones basadas en principios simplistas. Nos capacita para hacer juicios intuitivos sobre peligro, atracción y carácter. Estos juicios son rápidos e imperiosos. Algunos teóricos han supuesto, por tanto, que estamos equipados con una caja de herramientas innata de “heurísticos rápidos y frugales”. Y otros han supuesto incluso que no tenemos otra forma de razonar. Pero no deberíamos estar demasiado impresionados por el pensamiento inconsciente. También es el origen del pensamiento mágico.

El razonamiento inconsciente nos capacita para inferir relaciones causales basadas en restricciones, como por ejemplo: causas similares producen efectos similares. Vemos algo amenazador, y casi antes de haber sido conscientes de ello, los procesadores de bajo nivel nos están preparando para la defensa. Lo que varía de una inferencia inconsciente a otra no es el proceso, sino su contenido.

Las generalizaciones en las que se apoya el sistema pueden ser un resultado de la experiencia. Podemos adquirirlas sin saber qué son. Volviendo a mi ejemplo anterior, podemos aprender a reconocer un estilo musical sin haber sido nunca conscientes de los principios que utilizamos.

Las inferencias inconscientes pueden sesgar nuestra conducta sin que seamos conscientes de ellas. De la misma manera, podemos hacer una inferencia sin conocer en qué se basa. Fue una rápida conjetura intuitiva. Las intuiciones son útiles, y nos intrigan porque parecen no venir de ninguna parte y aun así ser útiles. Imaginemos que usted está intentando elegir entre, pongamos por caso, un conjunto de cuatro coches. Puede acceder a las descripciones de diversos aspectos de los coches, como su consumo, antigüedad y otras características similares. El razonamiento consciente es bueno mientras no haya demasiadas características a tener en cuenta. Pero si usted tiene información sobre una docena de características de cada coche, le supondrá mucho tiempo hacer una elección acertada. Ciertamente, un estudio ha mostrado que se tiende a hacer una elección más sensata cuando se evita pensar –al menos si se compara con la que se hace si solo se dispone de unos minutos para pensar–. Encuestas hechas a compradores mostraron que se consideraban más satisfechos con su elección si habían pensado cuidadosamente antes de comprar mercancías para las que son importantes pocas características, como manoplas de horno. Pero para aquellas compras para las que son importantes muchas características, como los ordenadores, estaban más satisfechos cuanto menos habían pensado sobre ellas.

Este tipo de estudios se prestan a ser mal interpretados. Los expertos pueden tomar decisiones intuitivas en un abrir y cerrar de ojos [blink] –para usar el título del provocativo libro de Malcolm Gladwell sobre el tema–. Los procesos inconscientes pueden tener en cuenta una gran cantidad de información debido a que se llevan a cabo en paralelo. Pero, como he sostenido, los procedimientos que esos procesos llevan a cabo son triviales desde el punto de vista computacional porque no tienen acceso a la memoria de trabajo. El razonamiento consciente puede llevar a cabo procesos más potentes, pero no puede tener en cuenta grandes cantidades de información. Se satura con facilidad. En ese caso uno puede desear desconectar sus deliberaciones conscientes. Si es un tipo de tarea en el que no hay una única respuesta correcta y es suficiente con una aproximación, puede no suponer un problema el dejar la mente en blanco –para usar el título de la parodia de Blink escrita por Noah Tall, y en palabras suyas explotar el poder de no pensar en nada en absoluto–. Pero, cuidado, las inferencias inconscientes son totalmente inapropiadas para muchas

tareas. No bastan para dirigir una central nuclear o para resolver el tipo de tareas de las que hablaré en otros capítulos de este libro, como por ejemplo descifrar el código Enigma nazi en la Segunda Guerra Mundial. Estos problemas exigen el razonamiento consciente, y por ello los razonadores expertos aprenden a tratar sus intuiciones con cierto escepticismo.

Notas al pie

[1] Nota de la trad.: En inglés la construcción del futuro se hace colocando delante del verbo un indicador de futuro [will].

[2] Nota de la trad.: Fenómeno que consiste en el desencadenamiento de determinados recuerdos a partir de un determinado indicio y puede producirse de manera inconsciente.

[3] Nota de la trad.: Examen que se debe realizar en muchas universidades norteamericanas para poder iniciar estudios de postgrado.

CAPÍTULO 6

Las emociones como inferencias

señales de interrogación, amistad, amenaza y

apaciguamiento, aceptadas instantáneamente, rara vez,

si alguna, malinterpretadas.

Auden

El razonamiento humano tiene una potencia limitada. Esta afirmación puede parecer extraordinaria a la luz de lo que ha conseguido –desde la comprensión profunda del mundo físico a la potencial solución de muchos de los problemas de la humanidad–. Sin embargo, es limitado si se compara con la inteligencia sobrehumana a la que he aludido en algunas ocasiones. Ciertamente, nos enfrentamos a problemas inferenciales que, debido a sus demandas computacionales, derrotarían incluso a ese poderoso ser. Como he señalado anteriormente, el razonamiento con múltiples premisas entre las que haya algunas con si, y y o es inabarcable. Otra fuente de inabarcabilidad es nuestra necesidad de compatibilizar múltiples objetivos y creencias, que no siempre son compatibles entre sí. Y otra fuente más es la necesidad de coordinar nuestras acciones con otra persona. Le envío a usted un correo electrónico invitándole a comer el martes próximo; usted me envía otro aceptando. Yo se lo devuelvo para que sepa que he recibido su aceptación; en caso contrario, usted podría pensar que su correo se ha perdido en el ciberespacio, y que no le esperaré para comer. Usted me responde para que yo sepa que usted sabe que he recibido su aceptación. Tal vez debiera escribirle para que sepa que yo sé que usted sabe que he recibido su aceptación. En realidad, solo aquellos que son puntillosos hasta la paranoia entran en esta serie interminable de correos electrónicos. Pero la tarea

de coordinación se vuelve aún peor si somos varios los que estamos intentando fijar una cita por este medio. Estos problemas pueden crecer hasta una dimensión que derrota a cualquier sistema computacional. Nuestro razonamiento tiene una potencia limitada.

Los problemas vitales que debemos resolver se reducen a los creados en nuestro entorno interno, como la fatiga y el dolor, los creados en nuestro entorno físico, como la necesidad de alimento y cobijo, y los creados en nuestro entorno social, como el deseo de una pareja y de descendencia. Nos las arreglamos y sin embargo en el laboratorio cometemos errores terribles. A veces los cometemos también en la vida real, aunque no es educado mencionar ese hecho en compañía de psicólogos que creen que en la vida diaria somos racionales. Esta discrepancia entre nuestro conocimiento y nuestra actuación es la paradoja fundamental de la racionalidad. De una parte, podemos entender los principios de la racionalidad –los cánones de la lógica y el cálculo de probabilidades–. De otra no conseguimos estar a su altura. La resolución de la paradoja es sencilla: el proceso de comprensión de los principios es abarcable, mientras que la tarea de seguirlos no lo es.

* * *

Si alguien nos insulta, por tanto, podemos tener una reacción emocional, como señaló Aristóteles. Entendemos el insulto y que amenaza nuestro autorrespeto (una evaluación cognitiva), se producen diversos cambios en nuestro cerebro (procesos neurofisiológicos) y, como resultado, las hormonas que se liberan en nuestro torrente sanguíneo nos preparan para la acción y nuestro corazón bombea más rápido (cambios somáticos), notamos cómo crece el enfado dentro de nosotros (un sentimiento subjetivo); hacemos un gesto de ira (una expresión facial) y decimos algo grosero como respuesta (una acción). Una típica experiencia emocional reúne, por tanto, pensamientos, sentimientos, cambios corporales y conductas. El simple hecho de saber cuáles de esos elementos son una “emoción” es motivo de discusión para los teóricos. Discuten también sobre las relaciones causales entre los diversos componentes. William James, por ejemplo, sostenía que no huimos porque estemos asustados: estamos asustados porque huimos. Las emociones, pensaba, surgen como una consecuencia de las

respuestas corporales. Sin embargo, esta visión inversa es contraria al sentido común y a las pruebas psicológicas. Sujetos totalmente paralizados siguen experimentando emociones.

Las emociones son un punto de encuentro entre la mente, el cuerpo y la conducta. En una emoción completa tenemos un sentimiento subjetivo, somos conscientes de cuál es su causa, experimentamos cambios corporales como consecuencia del funcionamiento de nuestro sistema nervioso autónomo y podemos generar varios gestos, expresiones y acciones. El sistema nervioso autónomo controla varios órganos, como por ejemplo las glándulas que liberan hormonas al torrente sanguíneo, y el sistema nervioso central controla nuestra conducta.

Las emociones son enigmáticas. Por ejemplo, ¿qué función cumplen? Algunos autores las consideran inútiles, como un órgano mental intrascendente que se inflama sin motivo lo mismo que el apéndice. Para muchas filosofías y religiones, las emociones deben ser evitadas –Platón, los estoicos, el místico ruso Gurdjieff, todos han arremetido contra ellas como disrupciones de la racionalidad–. Los budistas tibetanos no tienen un concepto de las emociones equivalente al nuestro, pero sostienen que el odio es destructivo, mientras que la verdadera iluminación produce una profunda felicidad y una afectuosa amabilidad. Jean- Paul Sartre, el filósofo existencialista, sostenía que cuando experimentamos una emoción, nos estamos ocupando de algo que no es más que pensamiento ilusorio. Sin embargo, el hecho de que todos los mamíferos sociales parezcan tener emociones sugiere que sí tienen una función.

Otro enigma es por qué somos conscientes de nuestros sentimientos. No somos conscientes de la liberación de hormonas o de otros cambios somáticos. No siempre somos conscientes de nuestras expresiones faciales, y no podemos reproducir de manera voluntaria una verdadera sonrisa, al menos una que pudiera engañar a un experto. Pero somos conscientes de que estamos contentos, tristes, enfadados o asustados. Incluso podemos ser conscientes de estos sentimientos sin saber qué los ha causado.

Quiero proponer una solución para estos enigmas, y mostrar cómo se afectan mutuamente las emociones y el razonamiento. Mi plan es seguir el curso de las emociones dentro de nuestra mente. Se crean mediante evaluaciones cognitivas, que pueden ser rudimentarias e inconscientes o complejas y conscientes. Podemos autopersuadirnos de una emoción. Incluso podemos autopersuadirnos

de una emoción referida a sucesos imaginarios. Sin embargo, una vez que sentimos una emoción, afecta a nuestro pensamiento. Por esta vía, estamos abiertos a lo que los ingenieros denominan “retroalimentación positiva”. Un determinado pensamiento produce una emoción, la emoción hace que pensemos algo más, el pensamiento aumenta la emoción, y con ello pensamos aún más, y así sucesivamente, hasta el punto de que el círculo puede convertirse en una espiral ascendente. Las emociones dominan nuestra vida mental y gobiernan lo que hacemos. En el capítulo siguiente continúo hablando de este tema.

* * *

Charles Darwin escribió en su clásico libro La expresión de las emociones en el hombre y los animales:

El poder de la comunicación entre los miembros de la misma tribu mediante el lenguaje ha tenido una importancia primordial en el desarrollo del hombre; y los movimientos expresivos de la cara y el cuerpo ayudan mucho a la fuerza del lenguaje.

Una teoría “comunicativa” formulada por Keith Oatley y yo toma como punto de partida este comentario, y postula que las emociones son comunicaciones entre miembros de la misma especie, en las que una señal que pertenece a un pequeño conjunto de señales mutuamente excluyentes expresa el estado emocional de un sujeto. También otra especie puede ser sensible a esas señales. Los perros, por ejemplo, parecen captar algunas emociones de sus dueños. Las señales que comunican las emociones son innatas, y por consiguiente universales en todas las culturas humanas: las expresiones faciales de la emoción son similares en todo el mundo. Su interpretación es muy diferente de la interpretación del lenguaje. El significado de una oración está compuesto de los significados de las partes que la componen de acuerdo con las relaciones gramaticales que existen entre ellas. Por esto es por lo que la interpretación de oraciones requiere memoria de trabajo. El significado de una expresión facial no está compuesto de

los significados independientes de lo que significan los labios, las cejas o las mejillas. Todos estos elementos forman parte de una señal inconfundible de una emoción, similar a las diversas señales de alarma, amenaza y sumisión que hacen otros mamíferos sociales. Su interpretación se lleva a cabo inconscientemente en regiones especializadas del cerebro. Si se reconstruyen fotografías de diversas emociones de forma que, pongamos por caso, los ojos pertenecen a una expresión de felicidad, pero los labios pertenecen a una expresión de enojo, etc., el resultado final es una quimera –una expresión extraña e ilegible–. La Figura 6.1 presenta algunas fotografías de expresiones faciales de emoción que encontré en la website de Paul Ekman, el investigador líder en el mundo sobre este tema. Las edité para generar quimeras. La Figura 6.2 presenta dos ejemplos y, como podrá usted ver, las expresiones son extrañas.

Fig. 6.1. Seis expresiones faciales de emociones (Copyright P. Ekman y W.V. Friesen, Consulting Psychologists Press, Palo Alto, CA).

Fig. 6.2. Dos expresiones faciales anómalas. La cara de la izquierda tiene los labios de la expresión feliz de la figura anterior impuestos en la cara enfadada de la misma figura, y la cara de la derecha tiene los labios de la expresión enfadada de la figura anterior sobre la cara feliz de dicha figura.

Para entender la siguiente suposición de la teoría comunicativa, necesitará usted recordar la arquitectura general de la mente, que describí al inicio de esta sección del libro. El cerebro tiene una estructura modular que tiene forma jerárquica, y la consciencia es similar a un sistema operativo que tiene un cierto control de los módulos, pero puede entrar en conflicto con ellos. Ciertos módulos pueden transmitir proposiciones con significado o modelos que las representan. La teoría comunicativa, sin embargo, postula una forma de comunicación interna alternativa, más antigua desde el punto de vista evolutivo y más rudimentaria: un pequeño conjunto de señales mutuamente excluyentes que se propagan de un módulo a otro. Estas señales son sencillas. Como las expresiones faciales, no necesitan de la memoria de trabajo para su interpretación, porque sus significados no están compuestos a partir de los significados de las partes que las componen. Son instintivas e inconfundibles. Un conjunto de señales de este tipo son las sensaciones corporales, como por ejemplo el hambre y la sed. Reflejan el control del medio interno. Otro conjunto de señales de este tipo son las emociones básicas. Circulan rápidamente a través del cerebro para dirigir la atención, movilizar los recursos corporales innatos y preparar una serie adecuada de conductas. La señal se inicia con una evaluación cognitiva, que puede producir o no un mensaje consciente, y un eslabón clave en la cadena es una transición inconsciente a la señal emocional. (La transición es una transición de Markov del tipo descrito en el capítulo anterior.). Por tanto, nuestras emociones son un tipo primitivo de razonamiento inconsciente que no produce conclusiones, sino señales.

Las señales emocionales nos ayudan a coordinar nuestros muchos objetivos y planes, dadas las restricciones de la presión temporal y de nuestros limitados recursos intelectuales. Pueden hacerse patentes en los momentos claves durante la ejecución de los planes. Son un sistema de control más flexible que los reflejos innatos, como por ejemplo los parpadeos. Pero son más rápidas que el razonamiento consciente porque no requieren utilizar la memoria de trabajo.

La experimentación de la emoción depende de que la señal emocional y el mensaje cognitivo se hagan conscientes. Si la señal se hace consciente, crea una sensación emocional característica. Esta sensación es común a cualquier tipo de situación que da lugar a la emoción. Si estamos colgados de una cuerda en la ladera de una montaña, a punto de hablar en público por primera vez o enfrentándonos a una persona violenta sentimos miedo –y la sensación, si la separamos de la evaluación cognitiva, es muy similar en todos esos casos, con la única diferencia de su intensidad–. Si el mensaje se vuelve consciente, sabemos la causa de nuestra emoción. Habitualmente, somos conscientes tanto de la señal como del mensaje. Pero de lo que no podemos ser conscientes es de la transición que está creando la emoción. Por tanto, la introspección nos puede decir, pongamos por caso, que estamos enfadados porque alguien nos ha insultado, pero no puede revelar la transición desde la evaluación hasta la emoción. Esta es la razón de que nuestras reacciones emocionales hacia la música instrumental sean tan misteriosas. Sabemos que la música ha causado la emoción, pero no entendemos por qué. La Figura 6.3 es un diagrama que resume la organización del sistema emocional en función de los dos canales de comunicación.

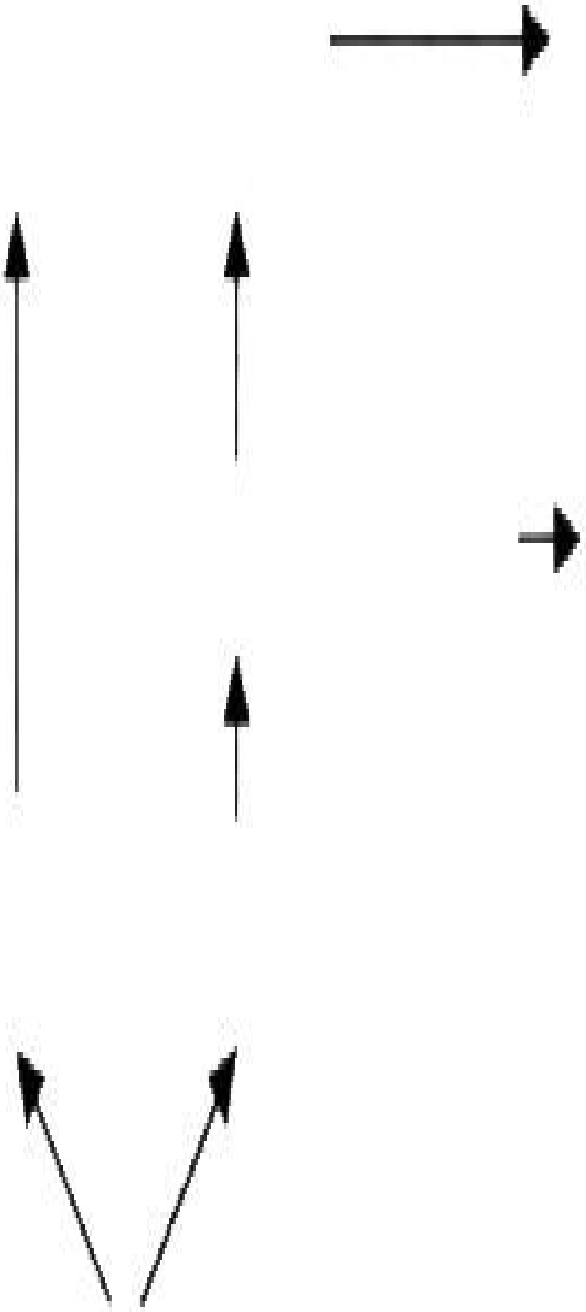


Fig. 6.3. Un diagrama esquemático de los dos canales de comunicación en la emoción.

La teoría comunicativa tiene en cuenta disociaciones entre la señal emocional y el mensaje cognitivo. La emoción puede ser inconsciente aun cuando el mensaje cognitivo afecte a nuestra consciencia. Por tanto, podemos no ser conscientes de nuestro estado emocional completo, y podemos incluso experimentar una "paralización" emocional o una falta de sensaciones casi patológica. Usted podría suponer que esta doctrina es freudiana, pero no lo es. Freud se salió de su camino para negar que las emociones pudieran ser inconscientes. En el caso opuesto, no somos conscientes de la evaluación cognitiva, y por tanto sentimos una emoción, pero no sabemos su causa. Decimos, por ejemplo:

Me siento feliz y no sé por qué.

Las diferentes emociones se sienten diferentes incluso cuando no sabemos sus causas.

La teoría comunicativa distingue entre emociones básicas y emociones complejas. Las emociones básicas son innatas y tienen sus propias señales inconfundibles en el cerebro y en las expresiones faciales universales. Incluyen la felicidad, la tristeza, el enojo, el miedo y la repugnancia. La amígdala, dos pequeños órganos a ambos lados del cerebro bajo la corteza, a la que está conectada, es crucial para las emociones básicas. Activa la liberación de hormonas, la supresión del dolor y diversas respuestas en el sistema nervioso autónomo que preparan al cuerpo para las emergencias. El daño en la amígdala afecta a las emociones básicas en las ratas, en los monos y en los humanos. También afecta a nuestra capacidad para reconocer las expresiones faciales de las emociones.

La predicción de que existen dos canales de comunicación independientes –una señal rápida de una emoción y un mensaje cognitivo más lento– se ha

comprobado en el caso de la emoción básica del miedo. Joseph LeDoux, un conocido neurocientífico, mostró que hay dos vías independientes por las que la información perceptiva llega a la amígdala. Una ruta es rápida. Va directamente del tálamo sensorial –un apeadero para la información perceptiva– a la amígdala. Evita el córtex, y por tanto depende de un análisis cognitivo rudimentario que puede dar lugar a la señal. La otra ruta es más lenta. Va del tálamo sensorial al córtex y luego a la amígdala. Es la ruta que sigue el mensaje sobre la evaluación cognitiva. La independencia relativa entre las señales y los mensajes se ha confirmado con observaciones clínicas. Una enfermedad conocida como "encefalitis letárgica" provoca que los pacientes sepan qué emociones deberían estar experimentando, pero no las sientan. Funciona el sistema de mensajes pero no el de señales. Otra enfermedad conocida como "alexitimia" hace que los pacientes tengan sensaciones, pero no sean capaces de articularlas. El sistema de señales funciona pero el sistema de mensajes no.

Si las emociones tienen una función comunicativa, las emociones básicas deben representar una variedad de situaciones posibles mediante un número reducido de señales diferentes, que se relacionen con entidades significativas en la vida de la especie. La teoría postula que el conjunto de emociones básicas, sus consecuencias somáticas y sus equivalentes expresivos son innatos e universales. Se relacionan con los objetivos de los seres humanos, que resumo en la Tabla 6.1 junto con ejemplos de las sensaciones y emociones que provocan. Las emociones son universales, pero la cultura influye en las evaluaciones cognitivas que las hacen aparecer.

Tabla 6.1. Los objetivos de los seres humanos y ejemplos de las sensaciones corporales y las emociones que crean

• La comida, la bebida, la respiración y la eliminación de desechos: hambre, sed, dificultad al respirar y sensaciones que surgen de la necesidad de orinar, defecar y exhalar.

• El mantenimiento de la salud: sensaciones de bienestar corporal o debilidad.

• La evitación de sustancias nocivas: emoción de asco.

• El mantenimiento de la vida y el logro de los objetivos: emoción de felicidad o frustración.

• La evitación de situaciones peligrosas: emoción de miedo.

• El mantenimiento de las relaciones entre los cuidadores y los niños: emoción de apego.

• Relaciones sexuales: sensación de deseo o frustración sexual. Emoción de amor o pérdida.

• Relaciones sociales con los iguales, incluidos superiores, iguales e inferiores: emoción de amistad u odio. Emoción de enojo o miedo.

Algunas emociones básicas se pueden experimentar por una razón desconocida, p. ej., la felicidad y la tristeza, y el enojo y el miedo. Otras emociones básicas tienen por objeto algo conocido. Se incluyen entre estas el amor sexual, la repugnancia, el odio, la emoción de apego de los niños a sus cuidadores, y la emoción recíproca de cuidado de los hijos por los padres. La señal emocional que se propaga en el cerebro depende del enlace de la emoción con la representación de un objeto. No podemos experimentar asco, pongamos por

caso, sin un objeto al que referir esa emoción.

Las inferencias inconscientes que subyacen en el pensamiento mágico, que discutí con anterioridad, hacen transiciones a emociones. Paul Rozin, el máximo experto en repugnancia, ha demostrado sus efectos. Supongamos que le ofrezco una bebida en una botella para muestras de orina nueva y sin usar. ¿La bebería usted? Probablemente no; la mayoría de las personas no lo haría. La función de la botella es "contagiosa" y la contamina. ¿Le apetecería una agradable mezcla de mantequilla de cacahuete y queso? Solo hay una pega: está en forma de cagada de perro. Los niños de uno a dos años, para horror de sus padres, disfrutan comiendo la mezcla. Los padres rechazan comerla, y puede que usted también lo hiciera. La repugnancia tiene su raíz en una aversión por el sabor desagradable de las sustancias tóxicas, pero sus causas son cognitivas. Lo que para un hombre es bueno para otro es veneno, como descubrió Darwin durante su viaje alrededor del globo. Estaba en Tierra del Fuego sentado tomando su almuerzo en una mesa portátil rodeado por un grupo de fascinados habitantes locales. Nunca habían visto carne en conserva, que era lo que él estaba comiendo. Uno de ellos metió su pulgar en la carne, y retrocedió, le repugnaba su húmeda esponjosidad. Darwin retrocedió también, le repugnaba la aparición del pulgar de un extraño en su comida. La dieta es un filón de pruebas que confirman las causas cognitivas de la repugnancia. Los budistas no comen carne; los hindúes no comen vaca; los judíos y los musulmanes no comen cerdo. Lo que comen los creyentes de una religión, los de otra lo rechazan. Y solo los ingleses adoran la marmite[4]. Las evaluaciones conscientes de naturaleza moral o religiosa pueden llevar a reacciones emocionales basadas en transiciones inconscientes. Yo creo: está mal comer perro, y por tanto si viera a alguien comiendo perro, me daría asco.

* * *

Las emociones básicas son la base biológica de las emociones complejas que parecen ser exclusivas de los humanos. Las emociones complejas dependen de inferencias conscientes que se relacionan con los modelos que tenemos de nosotros mismos y a menudo con comparaciones entre posibilidades alternativas o entre sucesos reales y posibilidades que imaginamos en historias alternativas.

Por tanto, solo pueden ser experimentadas por razones conocidas. Incluyen las emociones de remordimiento, indignación, orgullo y bochorno. Sentimos remordimiento, que es tristeza en relación con una acción o una falta de acción porque consideramos que hemos violado el código moral plasmado en el modelo idealizado que tenemos de nosotros mismos. Así que las emociones complejas integran una señal emocional básica y una evaluación cognitiva consciente. La concepción occidental del amor romántico también es una emoción compleja. Sus componentes innatos son el amor sexual y la felicidad. Sus componentes cognitivos incluyen el altruismo, y modelos idealizados de nosotros mismos y de aquellos a los que amamos. Estos componentes existen en muchas sociedades, pero su integración en un combinado reconocible y al que se le puede dar nombre fue un logro cultural.

Las emociones complejas parecen depender de una región de los lóbulos prefrontales del cerebro. Si esta región se daña, ya no experimentamos esas emociones, y dejamos de ser capaces de planificar nuestras vidas o de tomar decisiones sensatas. El caso clásico se refiere a un paciente llamado Phineas Gage, que a finales del siglo diecinueve trabajaba en la construcción de ferrocarriles en Estados Unidos. Estaba comprimiendo una carga explosiva con una barra de metal cuando la carga explotó y lanzó por el aire la barra que le atravesó la cabeza. Sobrevivió, pero sufrió un daño masivo en los lóbulos prefrontales. Como consecuencia de ello cambió totalmente y fue incapaz de desenvolverse en la vida. El neurocientífico Antonio Damasio ha descrito un paciente moderno similar, tenía éxito en los negocios y era un buen padre de familia hasta que desarrolló un tumor cerebral. La eliminación quirúrgica del tumor condujo a la extirpación de parte del tejido de los lóbulos prefrontales. Después de recuperarse de la operación, parecía normal en la mayoría de los aspectos. Seguía siendo muy inteligente. Experimentaba emociones básicas como por ejemplo el miedo. Las pruebas respecto a su memoria demostraron que era mejor de lo esperado para alguien con daño prefrontal. Sin embargo, era incapaz de hacer planes ni siquiera a unas horas vista. No podía desenvolverse en las relaciones sociales. Le echaron de los trabajos; se casó y divorció varias veces. Se arruinó. El papel del córtex prefrontal en el control de las emociones y de las relaciones sociales también se ha demostrado en monos. Su vida social está organizada en torno a “jerarquías dominantes” en la que los monos de un estatus más bajo rinden pleitesía a los monos de estatus superiores. Cada mono conoce su lugar en la jerarquía, del mismo modo que cada hombre inglés conocía su posición en la sociedad victoriana. Pero los monos con daño en los lóbulos prefrontales ya no siguen desempeñando su papel en la jerarquía.

Los shocks emocionales nos hacen sudar –una reacción en la que media el sistema nervioso autónomo–. El aumento de sudor en las palmas de las manos reduce la resistencia eléctrica de la piel. Por tanto, midiendo la resistencia en las palmas se detecta el cambio cuando vemos una escena que produce un shock. Esta reacción se conoce como la "respuesta galvánica de la piel", o RGP [GSR] en abreviatura. En los pacientes con daño de los lóbulos prefrontales no se da esta reacción. Tampoco son capaces de tomar decisiones sensatas en un juego experimental de apuestas que Damasio diseñó para imitar las decisiones de la vida cotidiana. Tenían que elegir cartas de diversos montones, y dependiendo de cuál fuera la carta, perdían o ganaban una cierta cantidad de dinero simulado. Dos de los montones producían de forma sistemática recompensas pequeñas. Otros dos producían recompensas bastante mayores pero ocasionalmente una pérdida masiva. Los participantes sin daño cerebral desarrollaron rápidamente una respuesta RGP anticipándose a estas pérdidas, y aprendieron a rechazar esos montones –incluso antes de que pudieran explicar por qué–. Pero los pacientes con daño prefrontal nunca aprendieron a evitarlos. Lo mismo que en la vida, en el juego también se arruinaron.

La teoría comunicativa postula que las emociones aseguran el que tomemos decisiones en lugar de perdernos en un laberinto de razonamiento. La transición a una emoción proporciona una señal, o de buenos resultados que nos animan a perseverar, o de malos resultados que nos previenen de un peligro inminente. Estas reacciones viscerales, o "marcadores somáticos" como los llama Damasio, nos capacitan para responder con el mínimo pensamiento.

* * *

La evaluación cognitiva que conduce a una emoción puede ser consciente o inconsciente, pero la transición a la emoción es siempre inconsciente. No podemos conectarla o desconectarla. Como decía Aristóteles, no podemos elegir sentir enojo o miedo. Nuestras sensaciones son involuntarias. Para San Agustín, una de las consecuencias del pecado de Adán fue que nuestra carne desobediente se situó por encima de nuestra voluntad. Antes de la caída, el hombre podía conseguir erecciones a voluntad; a partir de entonces, las erecciones podían ser involuntarias, insistentes, inoportunas. Al contrario que los abogados, Auden

decía que no había problema para reconocer la pornografía –pida simplemente a un jurado que vote, y vea cuántos miembros masculinos se yerguen–. Tampoco podemos controlar nuestras emociones. Vemos a una niña llorando porque se ha quedado huérfana y sentimos una profunda simpatía por ella. Vemos una muestra de crueldad bárbara y sentimos indignación. No importa cuánto esfuerzo hagamos, no podemos ejercer un control directo sobre nuestras sensaciones. Como máximo, podemos distraernos tratando de pensar en otra cosa, o podemos intentar reevaluar los hechos que causan las emociones. No podemos estar felices (o tristes) como resultado de una simple intención voluntaria. La única forma en que podemos obtener una emoción es pensando en algún episodio que sabemos que es probable que la provoque. Podemos ser capaces de controlar la expresión de nuestras sensaciones. Pero, incluso con una fuerza de voluntad de hierro, nuestros verdaderos sentimientos pueden filtrarse al exterior en forma de breves expresiones faciales. Como ya mencioné, una filmación de nuestro rostro a cámara lenta puede revelar un involuntario “filtrado”. Podemos estar sonriendo para enmascarar nuestro enojo pero, por un instante, de manera apenas visible, fruncimos el ceño.

Nuestras emociones con frecuencia nos empujan a actuar. Pero seguimos pudiendo razonar, y algunas veces nuestro razonamiento entra en conflicto con nuestra emoción. Podemos inferir que una acción sería inmoral o imprudente. La resolución del conflicto es una cuestión abierta: no podemos predecirla. En nuestra conducta existe un elemento irreducible de no determinismo. Algunas circunstancias llevan a fuertes sensaciones que son inadecuadas. Hay a quien, por ejemplo, le aterrrorizan las arañas. Otros son superados por el pánico a entrar en un ascensor. Estas emociones son anormales, y las tomaré en consideración más adelante.

Las novelas, las obras de teatro y las películas pueden todas ellas despertar emociones reales sobre hechos irreales. Lloramos o reímos por algo que sabemos que es ficticio. La música es aún más enigmática, porque nos puede conmover aun cuando no se refiera a nada en absoluto. Nos sentimos felices como consecuencia de escuchar una pieza de música, pero la música no es el objeto de nuestra emoción, sino su causa. No podemos decir por qué esa pieza de música nos alegra mientras aquella nos entristece. La razón, como he sostenido en otro lugar, es que la música imita conductas emocionales, y que las evaluaciones cognitivas que dan lugar a las emociones básicas son rudimentarias. No valoran la realidad de los sucesos, sino que dependen de transiciones inconscientes. Esta tendencia a experimentar emociones reales frente a sucesos irreales parece ser

universal en todas las culturas.

* * *

El tráfico entre razonamiento y emociones se mueve en ambas direcciones. Las inferencias evocan emociones: las emociones evocan inferencias. Aun cuando la conexión principal en la evocación de una emoción sea una transición inconsciente a partir de una evaluación, el primer hecho en la cadena causal puede ser una inferencia consciente. Extraemos una conclusión de la que somos conscientes, y eso nos lleva a una transición inconsciente a una emoción. Hace unos años, salí un día de mi oficina, fui a donde había aparcado el coche, y vi que había desaparecido. Alguien me lo ha robado (inferí), y por ello me sorprendí y me enfadé. Volví a la oficina y llamé a la policía para informar del robo. Salí otra vez, y giré por una calle diferente para coger el autobús hasta casa. Cuando caminaba por la calle, vi mi coche aparcado allí cerca. Lo había aparcado allí por la mañana, porque la zona donde aparcaba habitualmente estaba llena. Mi enojo fue sustituido inmediatamente por alivio, pero también por bochorno por mi mala memoria y mi estupidez por inferir que me habían robado el coche. Esa inferencia me había hecho enfadar, y quizá la emoción había nublado mi memoria.

Las inferencias que crean emociones pueden referirse a sucesos imaginarios o hipotéticos. Leemos en el periódico que se ha producido un incendio en un piso cerca de donde nosotros vivimos. Imaginamos que podía haber ocurrido en nuestro edificio, y experimentamos un tufillo de ansiedad. Nuestra capacidad para imaginarnos en la piel de otros e incluso para imaginar lo que están pensando es la base de la empatía. Por consiguiente, toda la gama de inferencias que van desde las intuiciones inconscientes a las deducciones conscientes puede crear emociones.

Cuando razonamos sobre nuestras emociones futuras, a menudo nos equivocamos. Imagínese que hereda una gran cantidad de dinero, ¿será usted feliz para siempre? Puede pensar que sí. Imagine que en lugar de eso se ve implicado en un accidente de consecuencias desastrosas que le deja tetrapléjico. La vida no merecerá la pena, ¿no? Está equivocado. Ninguna de las predicciones

es correcta. Después de algunos años, los ganadores de grandes premios de lotería no son, por término medio, más felices que el resto de nosotros, y los tetrapléjicos no son menos felices que los demás. Así es como ellos se valoran en una escala de felicidad. Puede usted decir: quizá están equivocados. Pero, ¿de qué otra forma podemos evaluar sus sensaciones? Si sigue siendo usted escéptico, piense en las consecuencias de un cambio en su vida –sea a mejor o a peor– y compare cómo se siente ahora respecto a él y cómo anticipó que se sentiría. Las emociones, como las sustancias radiactivas, tienen una vida media. Se desintegran con el tiempo.

* * *

Nuestras emociones nos afectan cuando elegimos entre varias opciones. Para los economistas clásicos es racional elegir cualquier opción que produzca la mayor "utilidad esperada". Utilidad se refiere al valor de un resultado. Un regalo de cien dólares tiene una cierta utilidad y también el que te metan un palo en un ojo –una utilidad bastante negativa–. Los economistas suponen que todas las opciones pueden colocarse en una única escala de utilidades. Esperada se refiere a lo probable de un resultado, y por tanto para computar la utilidad esperada, multiplicamos nuestra estimación de probabilidad de un resultado por la utilidad que tiene para nosotros, p. ej., un regalo instantáneo de 95 dólares tiene la misma utilidad esperada que un 95 % de probabilidades de conseguir 100 dólares.

Mi antiguo colega de Princeton Danny Kahneman y el difunto Amos Tversky mostraron que violamos los principios de racionalidad basados en las utilidades esperadas. Durante mucho tiempo, muchos economistas rehusaron creer en sus resultados, pero las pruebas se impusieron y Kahneman ganó el premio Nobel de economía en 2002. He aquí un ejemplo de una observación típica que puede comprobar usted mismo. Cuál de estas opciones elegiría:

¿Una ganancia segura de 900 dólares o un 90 por ciento de probabilidades de ganar 1.000 dólares?

Como la mayoría de nosotros, probablemente usted se decide por la ganancia segura. ¿Por qué arriesgarse a no ganar nada por tener la oportunidad de un mísero extra de 100 dólares? Ahora, en cambio, suponga que tiene usted que sufrir una pérdida. ¿Cuál de estas dos opciones preferiría:

¿Una pérdida segura de 900 dólares o una probabilidad de un 90 por ciento de perder 1.000 dólares?

Como la mayoría de nosotros, probablemente usted preferiría la segunda opción. Puede usted perder 1.000 dólares, pero existe una oportunidad de librarse sin perder un céntimo. Merece la pena arriesgar los 100 dólares extra por tener la oportunidad de no perder nada. Si tuviéramos una serie de ganancias como las de la primera situación, la cantidad que esperaríamos ganar sería la misma para ambas opciones: si ganamos 1.000 dólares en el 90 por ciento de las ocasiones, nuestras ganancias medias totales son de unos 900 dólares. Así que, ambas opciones tienen la misma utilidad esperada positiva. De forma similar, una serie de pérdidas como las de la segunda situación tienen la misma utilidad esperada negativa (una pérdida de 900 dólares). Por tanto, con ganancias que tienen la misma utilidad esperada, rechazamos la apuesta en favor de lo seguro; pero con pérdidas que tienen la misma utilidad esperada, rechazamos lo seguro en favor de la apuesta. ¿Por qué?

La utilidad esperada no puede explicar la diferencia, pero nuestras emociones sí. La teoría de las “perspectivas” [“prospect” theory] de Kahneman y Tversky refleja una asimetría en el valor que asignamos a las ganancias y a las pérdidas, y postula que evaluamos estos valores en relación con el status quo o con algún otro punto de referencia relevante. El dolor por una pérdida es proporcionalmente mayor que el placer de una ganancia equivalente, y por eso tendemos a correr riesgos para evitar pérdidas, pero no para conseguir ganancias.

Para la teoría clásica, lo único que importa cuando elegimos es la utilidad esperada. Sin embargo, cuando elegimos, más que calcular utilidades, razonamos. Inferimos cómo nos sentiremos como consecuencia de los posibles resultados. Por ejemplo, ¿vacunaría usted a su hijo contra la tosferina cuando existe una mínima oportunidad de que la vacunación tenga un efecto secundario

mortal? Mucha gente rehúsa hacerlo. Anticipan lo terriblemente mal que se sentirían si su hijo muere, aun cuando su hijo corre un riesgo mayor de morir de tosferina. La evaluación del riesgo no es fácil, y el terror a consecuencias fatales pesa sobre nosotros.

* * *

Nuestras emociones nos preocupan. Como resultado, podemos equivocarnos al razonar. El estrés y la ansiedad son causas de fallos cognitivos conocidas desde hace tiempo, desde pilotos que apagan el motor equivocado cuando hay un motor ardiendo, hasta los ingenieros de Chernobil, que no quisieron creer que se estaba destruyendo el reactor. Las emociones parecen impedirnos pensar en posibilidades alternativas. En el laboratorio se producen efectos similares. Si vemos una película que inspira una emoción de felicidad o tristeza, es menos probable que hagamos las elecciones correctas en la tarea de selección de Wason. Recordará usted que en esta tarea teníamos que seleccionar pruebas para determinar la verdad o falsedad de una afirmación general. Lo hacemos mejor si la película no es emotiva o si no vemos ninguna película.

Podría usted sospechar que cuando estamos contentos razonamos mejor que cuando estamos tristes. Docenas de experimentos apoyan esta suposición. Una serie de estudios pedía a los participantes que resolvieran problemas que dependían de un insight clave. Un ejemplo de problema de este tipo, conocido hace tiempo por los psicólogos, es el problema de la “vela”. La tarea consiste en fijar una vela en una puerta disponiendo solo de la vela y una caja de chinchetas. Podemos intentarlo atravesando la vela con una chincheta, pero no es lo bastante larga como para conseguirlo. El insight que resulta crucial es comprender que la caja que contiene las chinchetas también puede servir como sustentación de la vela. Utilizamos las chinchetas para clavarla a la puerta, y luego ella sostiene la vela. Darse cuenta de que un objeto que desempeña una función puede desempeñar otra parece fácil, pero no lo es. En realidad, es tan difícil que los psicólogos tienen en su jerga un término para ese problema: lo llaman “fijeza funcional”. Era más probable que llegaran a un insight crucial los participantes que se sentían felices porque habían visto un corte de una película cómica. Sufrían menos fijeza funcional que aquellos que habían visto una película triste o

neutra, o no habían visto ninguna película. La felicidad ayuda al insight.

Una emoción nos puede preocupar como nos preocupa que un hombre nos golpee en la cabeza con un mazo de goma, o puede apresurarnos o retrasarnos dependiendo de si nos sentimos optimistas o pesimistas. Pero una emoción puede ser adecuada para la tarea que estamos realizando. Los experimentos psicológicos han establecido una serie de resultados de este tipo: La ansiedad nos lleva a pensar en riesgos futuros, y nos consideramos en mayor riesgo que cuando la valoración la hacemos sin sentir ansiedad. La ansiedad también nos hace estar más vigilantes cuando hay varios hechos que compiten por nuestra atención. La depresión, por el contrario, no lo hace, sino que nos centra en el pasado: es una emoción de pérdida, y cavilamos sobre acontecimientos pasados. Aumenta nuestro recuerdo de sucesos tristes, mientras que la ansiedad no afecta a nuestro recuerdo de sucesos amenazantes. Asimismo, la depresión leve hace que seamos más exactos al calcular nuestro control respecto a los resultados en los que ganamos o perdimos pequeñas sumas de dinero. Estamos más tristes pero somos más sensatos.

Tiene usted ahora diez segundos para responder a esta pregunta:

Soy peor que mi jefe.

Mi jefe es peor que mis colegas.

¿Quién es el peor?

Para la mayoría de nosotros, este problema es más difícil que otro que no sea deprimente. No pensamos a menudo en los aspectos negativos de nuestra propia actuación. Pero, si estamos deprimidos, estamos acostumbrados a cavilar sobre nuestras deficiencias –quizá demasiado acostumbrados a ello– y por tanto el problema anterior no nos resulta tan difícil como uno sobre nuestra felicidad.

Keith Oatley y sus colegas han demostrado efectos similares en las inferencias que los participantes hacen respecto a historias. Por ejemplo, en la historia corta de Russell Banks, Sarah Cole: a type of love story, un hombre sin corazón rompe una relación amorosa con gran crueldad. El efecto sobre los participantes en un experimento fue entristecerles o enfadarles. Los que estaban tristes tendían a razonar sobre los sucesos iniciales de la historia. Los que estaban enfadados tendían a razonar sobre sucesos futuros. Como escribió Aristóteles, el enojo nos lleva a razonar sobre formas de venganza estricta.

Las emociones nos hacen razonar sobre sus causas. Para poner un ejemplo inspirado en Kahneman y Tversky, supongamos que usted conduce en dirección a un aeropuerto para coger un avión y en el camino se para a beber algo. Llega al aeropuerto unos minutos tarde, y pierde el avión por poco. Está decepcionado, y su emoción tiende a hacerle pensar para usted mismo:

Si no me hubiera parado a tomar algo, no habría perdido el vuelo.

El pensamiento es “contrafactual” porque se refiere a una alternativa imaginaria a los hechos. Se refiere a dos situaciones: qué pasó y qué podría haber pasado en la situación alternativa, pero contrafactual, en la que usted no hubiera parado a tomar algo. En un libro brillante, The Rational Imagination, Ruth Byrne mostró que se necesita imaginación para pensar sobre posibilidades contrafactuales. Las tenemos en cuenta para aprender de nuestros errores, para corregir planes que han fallado, y a veces para alegrarnos de nuestra suerte. Pero, cuando las cosas salen mal, pensamos a menudo en posibilidades contrafactuales. Muchas emociones complejas, como el arrepentimiento, la culpa y la vergüenza se suscitan de esa forma. Ellas a su vez nos predisponen a pensar en un filón contrafactual.

Cuando comparamos lo que ha sucedido con lo que podía haber sucedido, podemos imaginar mejores resultados. Y a veces nos consideramos heridos cuando pensamos en un resultado peor. He aquí una historia que lleva a mi condicional contrafactual favorito. Una vez conduje sin parar desde Italia hasta la costa norte francesa para llegar a mi casa en Inglaterra lo antes posible. Al llegar al puerto de Dieppe al final de la tarde, vi el ferry que zarpaba del muelle.

Lo había perdido por cinco minutos. Pregunté cuándo salía el siguiente ferry y me dijeron que a las 10 de la noche. Para consolarme, salí a tomar una magnífica cena francesa. Después de cenar, volví al puerto. Noté que estaba empezando a soplar una tormenta, y el hombre que controlaba la entrada del puerto me dijo que el ferry de las 10 había sido cancelado a causa del tiempo. No habría más ferrys hasta el día siguiente. Con una redoblada decepción, llamé a mi mujer y me registré en un hotel para pasar la noche. A la mañana siguiente, de nuevo en el muelle, llegó por fin un ferry. Los pasajeros del puente parecían desesperados, ligeramente "blancos como el papel", y lo que era aún más extraño, no desembarcaban. Era el mismo ferry que yo había perdido por poco el día anterior. El estado del mar era tan malo que no había podido entrar al puerto de Newhaven en Inglaterra, y los pasajeros habían pasado toda la noche dando vueltas por el barco durante una de las peores tormentas que se recordaban en el Canal de la Mancha. Si no hubiera perdido el barco, yo habría estado también con ellos.

* * *

Las emociones son tan importantes para nosotros que un número sorprendente de personas piensa que son exclusivas de los seres humanos. Un ordenador puede ser capaz de ganar al campeón mundial de ajedrez, dicen, pero no puede tener sensaciones. En realidad, las pruebas implican que las emociones son anteriores al Homo sapiens al menos en doscientos millones de años. Todos los mamíferos sociales, incluidas las ratas, tienen los órganos cerebrales de las emociones básicas y todos ellos parecen experimentar estas emociones.

La función de las emociones es prepararnos para diversos recorridos de acción o inacción. Nos hacen reaccionar más rápidamente que el razonamiento consciente. Nos hacen inteligentes en una emergencia, pero pueden atontarnos cuando inundan nuestro razonamiento consciente. Comunican el estado de nuestra mente a los demás. Normalmente somos conscientes de la causa de una emoción. Nunca somos conscientes de la transición a la emoción. Normalmente somos conscientes de la emoción, y ello puede influir en nuestras acciones intencionales. Sin embargo, cuando nuestro intelecto tira de nosotros en un sentido y nuestra emoción tira de nosotros en otro, somos incapaces de

desconectar la emoción. A veces ni siquiera podemos controlar nuestra conducta. Aquí tenemos la descripción de Darwin sobre el resultado de uno de estos conflictos:

Acerqué la cara a la gruesa placa de cristal que había frente a una víbora brumadora en el parque zoológico, con la firme determinación de no retroceder si la serpiente me atacaba; pero, tan pronto como lanzó el golpe, mi resolución no sirvió de nada, y salté una o dos yardas hacia atrás con asombrosa rapidez. Mi voluntad y mi razón fueron impotentes frente a la imaginación de un peligro que nunca había experimentado.

Notas al pie

[4] Nota de la trad.: Marmite es una pasta hecha a base de levadura que se come untada en las tostadas.

CAPÍTULO 7

El razonamiento en las enfermedades psicológicas

CAPÍTULO 7

El razonamiento en las enfermedades psicológicas

He aquí una persona razonando sobre un dolor de estómago:

Estoy asustado por el ligero dolor que siento en el abdomen en el lado del hígado. Podría ser un síntoma de cáncer, un cáncer de hígado. Recuerdo a un tío mío que murió de cáncer de hígado después de sufrir mucho. Pero tenía ochenta y tantos años y yo tengo treinta, y un cáncer de hígado a mi edad es raro. Por otra parte, no es imposible. Además, me parece que tengo aspecto enfermizo; tengo la lengua sucia; a veces tengo un sabor amargo en la boca. Parece que estoy pálido, y podría tener anemia. Por supuesto son síntomas comunes y pueden ser triviales. Los he tenido otras muchas veces. Pero están ahí y no son incompatibles con el cáncer. Es más, no lo excluyen. El médico me mandó varias pruebas, y todos los resultados fueron negativos. Pero los resultados podrían ser los de otra persona –a veces los laboratorios mezclan los tubos de las pruebas, o la secretaria comete un error al escribir el nombre del paciente, o pone los resultados de una persona en el sobre de otro paciente–. Siempre puede producirse un error. El laboratorio puede ser muy profesional, pero no puede haber un cien por cien de garantía de que no cometió un error. Además, yo soy el principal responsable de mi propia salud. Se puede imaginar cómo me sentiría si realmente tuviera cáncer y lo hubiera dejado para demasiado tarde. Lo mejor que puedo hacer es volver al médico.

Ahora, he aquí a una persona diferente razonando sobre el mismo tema:

Estoy asustado por el ligero dolor que siento en el abdomen en el lado del hígado. Podría ser un síntoma de cáncer, un cáncer de hígado. Recuerdo a un tío mío que murió de cáncer de hígado después de sufrir mucho. Al principio sus síntomas eran los mismos que los míos: tenía un dolor de estómago similar. No se preocupó, y los médicos le dijeron que no estaba enfermo. Pero, mientras tanto, el cáncer se estaba extendiendo. Ahora, el cáncer puede estarse extendiendo de la misma forma por mi abdomen. Además mis síntomas parecen haber empeorado durante las últimas semanas. Nadie me cree y nadie me toma en serio. Cuando empiecen a tratarme ¡será demasiado tarde! Además me parece que tengo aspecto enfermizo; tengo la lengua sucia; a veces tengo un sabor amargo en la boca. Parece que estoy pálido y podría tener anemia. ¡Qué trauma para mí y mi familia cuando el cáncer sea correctamente diagnosticado y sea demasiado tarde! Después tendré una vida de sufrimiento, medicamentos, pruebas médicas, revisiones y operaciones quirúrgicas. Lo mejor que puedo hacer es volver al médico.

La primera estampa es característica de una persona que sufre un transtorno “obsesivo compulsivo”. Cavila sobre una posible enfermedad, y razona de una forma dialéctica, oscilando entre argumentos a favor y en contra de la peor situación posible y nunca es totalmente capaz de descartar el segundo. La segunda estampa es característica de un paciente que sufre de hipocondría. Su razonamiento se inicia con los síntomas corporales, se centra en ellos y amplifica las nefastas consecuencias de no hacerles caso.

Este capítulo se ocupa de las enfermedades psicológicas y del razonamiento cuando estamos afectados por ellas. Sin embargo, en primer lugar necesito tomar en consideración una cuestión crucial. ¿Existen enfermedades que en lugar de estar causadas por una patología del cerebro tienen un origen psicológico? Algunos psiquiatras lo dudan. Pero, otros muchos creen en la aparición de “neurosis”, como se llamaba a esas enfermedades hasta que la American Psychiatric Assotiation eliminó la palabra de su manual de diagnóstico en 1980. Estas enfermedades, si existen, deben tener sus orígenes en experiencias que afectan a los pensamientos y sensaciones posteriores y que nos llevan a sufrir, a menudo hasta el punto de buscar ayuda profesional. Un organismo que no puede modificar su conducta como resultado de la experiencia –una ameba, tal vez– no

puede llegar a estar mentalmente enfermo. Podría mostrar síntomas de transtorno mental, pero la causa debería ser algún daño en su sistema nervioso.

La distinción entre los transtornos mentales y los que se originan por un daño en el cerebro parece dualista: existe la mente, existe el cerebro, y los dos interactúan de vez en cuando, quizá a través de la glándula pineal (como imaginó Descartes). Sin embargo, el cerebro podría ser la base material de la mente. Imagínese un robot controlado por un software complejo que se modifica a sí mismo como resultado de la experiencia. Aprende, pongamos por caso, que el robot no debe intentar pasar por huecos que son demasiado estrechos. Pero puede adquirir una conducta patológica, como no pasar por ningún hueco: ningún programa de aprendizaje puede ser infalible, debido a la inabarcabilidad de la inducción. El robot también podría sufrir daño en el hardware en el que corre su programa. Podemos por tanto establecer una distinción razonable entre fallos que son resultado de la experiencia, produciendo problemas en el programa, y fallos que son resultado de un daño en su unidad de procesamiento, produciendo daños más profundos[5]. Hacemos una distinción similar cuando decidimos que un ordenador ha funcionado mal debido a un fallo en su procesador de texto, o debido a un problema con su disco duro.

Aceptando la distinción entre mente y cerebro, podemos plantear la pregunta: ¿qué causa la enfermedad mental? Nuestro conocimiento sobre la respuesta es peor que, pongamos por caso, el conocimiento médico sobre el cólera hace ciento cincuenta años. Al menos los médicos se dieron cuenta entonces de que era una enfermedad epidémica que venía de Asia, y mataba aproximadamente a la tercera parte de los que se infectaban. Sobre lo que no estaban de acuerdo era sobre la causa de la enfermedad. Con las enfermedades mentales, sin embargo, no hay consenso ni sobre su existencia ni sobre su causa. La propensión humana a tratar las enfermedades no se ve limitada de ningún modo por la falta de conocimiento. Hay varios cientos de variedades de psicoterapia, desde la terapia analítica activa hasta la técnica psicoenergética Zaraleya.

Los psiquiatras que rechazan la enfermedad mental atribuyen la patología a un cerebro defectuoso. El desarrollo de medicamentos adecuados que alteren el cerebro, o la terapia génica, dicen, es la mejor forma de cura. Un enigma de gran importancia que se plantea ante esta opinión es el que la psicoterapia puede curar la enfermedad mental. Es más, algunos pacientes se recuperan sin tratamiento, aunque la psicoterapia lleva a una recuperación más rápida. Si la experiencia puede curar una enfermedad, entonces quizá la experiencia puede causar la

enfermedad.

Los conductistas sostienen que las enfermedades mentales son malos hábitos. Los pacientes aprendieron respuestas inadaptadas frente a ciertos estímulos. Por consiguiente, el tratamiento es enseñarles respuestas más apropiadas frente a esos estímulos. La terapia de conducta puede eliminar las respuestas inadaptadas. No obstante, el que sea una cura efectiva es motivo de disputa.

Entre los que admiten una causa mental de la enfermedad, Freud y sus seguidores y rivales suponen que los conflictos inconscientes emergen en la consciencia de una forma distorsionada que produce neurosis. Freud intentó establecer una psicología científica, pero ha sido difícil evaluar su teoría. Los psicólogos pueden evaluar fragmentos y partes de ella, pero no la hipótesis fundamental de Freud, que él sostuvo al final de su vida: los bebés tienen sensaciones sexuales, y su represión –un proceso inconsciente en sí mismo– causa neurosis. Los psicoanalistas apoyan esta hipótesis con las intuiciones que obtienen al tratar a los pacientes. Estas intuiciones pueden ser correctas, pero no son más seguras que las intuiciones de los doctores de comienzos del siglo diecinueve sobre el cólera. De lo que podemos estar seguros, sin embargo, es de que los procesos inconscientes se refieren a algo más que a anhelos sexuales inaceptables y deseos destructivos. Un amplio conjunto de procesos de los que no somos conscientes hace posible la consciencia.

La teoría psicoterapéutica sufrió una revolución hace treinta años con el desarrollo de la “terapia cognitiva”. Aaron Beck y el resto de los que la propusieron sostenían que la causa de las neurosis es el razonamiento defectuoso, sea por inferencias no válidas o por inferencias a partir de falsas creencias. Las fuentes de las falsas creencias son el testimonio de los otros, las percepciones inexactas y el razonamiento no válido. Pero es improbable que el testimonio de los otros sea un factor en la creación de falsas creencias que causen enfermedades psicológicas. De la misma forma, quienes padecen enfermedades psicológicas, a diferencia de psicosis más serias como la esquizofrenia, rara vez sufren fallos en la percepción. De ahí que, en la versión de Beck, la principal fuente de falsas creencias es el propio razonamiento defectuoso. Ciertamente, la idea es omnipresente. Incluso Auden, al que he citado con frecuencia en este libro, escribió: “El neurótico es alguien que extrae una conclusión general falsa a partir de un ejemplo particular.”

La terapia cognitiva se ha utilizado para tratar diversas enfermedades

psicológicas, incluidas la ansiedad y la depresión. Sin embargo, el principio fundamental de la teoría sobre el razonamiento defectuoso no tiene un apoyo firme. Mi propósito es por tanto reexaminar el papel del razonamiento en la psicopatología y, a la luz de los datos recientes, proponer una nueva teoría de las enfermedades psicológicas desarrollada junto con mis colegas italianos Francesco Mancini y Amelia Gangemi. Esta teoría, a la que nos referimos como la teoría de la "hiperemoción" para las enfermedades psicológicas, se apoya en la descripción del razonamiento inconsciente y las emociones que ya se ha presentado en esta sección del libro.

* * *

Lo que, según esta teoría, sucede en una enfermedad mental es que hacemos una evaluación cognitiva de una situación. Esto nos lleva mediante una transición inconsciente a una emoción básica que es adecuada para esta evaluación, pero que es desproporcionada en cuanto a su intensidad. Es una hiperemoción. Somos conscientes de la naturaleza anómala de nuestra respuesta, y por ello razonamos sobre ella. Este razonamiento, a su vez, puede ampliar nuestros problemas. Como ejemplo, consideremos el caso de la mujer, mencionada en la Introducción, que examinaba una fotografía de Rock Hudson en un periódico. Tenía una emoción básica de intensidad inadecuada: se preocupaba por haber contraído el VIH al haber tocado el periódico. Razonaba sobre sus preocupaciones, y el efecto era amplificar sus problemas.

¿Cuál es el origen de las transiciones inconscientes inadecuadas? Estas transiciones nos ocurren a todos, y por tanto podemos tener tendencia a las emociones inadecuadas y a los problemas mentales. En su forma más suave, no nos causan problemas. Conducen a pensamientos y, algunas veces, a acciones que son irracionales. Pero podemos ser felices e incluso tener éxito si, pongamos por caso, tenemos una fobia irracional durante toda nuestra vida. El entorno afecta. El estrés prolongado, como el que experimentaron los soldados en las trincheras durante la Primera Guerra Mundial, nos hace susceptibles de sufrir enfermedades psicológicas. También nos diferenciamos en nuestra susceptibilidad innata respecto a ellas. Los gemelos idénticos proceden de un mismo óvulo y tienen genes idénticos, si un gemelo tiende a ser neurótico el otro

también, y la correlación es mayor que entre gemelos que no son idénticos. Por tanto, existe una predisposición genética para la enfermedad mental, y la investigación está empezando a localizar los genes que resultan relevantes.

* * *

La transición inicial hacia una emoción anómala se produce como resultado de un suceso que podemos recordar a menos que sucediera en la infancia. El suceso puede ser inocuo para otros, pero para nosotros es traumático debido a la emoción consiguiente. Ahora bien, puede ser un trauma real. Por ejemplo, a una paciente le asustaba excesivamente el volar después del intenso miedo que había pasado en un avión que había sido alcanzado por un rayo. Los sucesos traumáticos pueden predisponer a los sujetos a tener estas reacciones. Un taxista de Nueva York me contó –mientras estábamos en un atasco en un pequeño túnel en Manhattan– que, después de ver el colapso de las torres del World Trade Center el 11 de septiembre de 2001, estaba a punto de entrar en el túnel Holland bajo el río Hudson cuando le invadió una intensa ansiedad. Le aterrorizaba poder quedarse allí atascado y que le diera un ataque de pánico. Había tardado varios años en recuperarse y ser capaz de conducir de nuevo por un túnel. Son pocos los pacientes incapaces de recordar los orígenes de sus enfermedades mentales –una afirmación confirmada por un estudio del que informaré posteriormente–. La nueva teoría permite, sin embargo, que, como resultado de fluctuaciones arbitrarias, se pueda producir en cualquiera una transición inconsciente a una hiperemoción. Los sucesos posteriores y nuestra constitución mental determinan si es solo una anomalía momentánea o se asienta de manera más profunda.

La teoría de la hiperemoción implica una taxonomía de las enfermedades mentales. Pueden surgir a partir de cualquiera de las sensaciones corporales o emociones básicas de la vida de los mamíferos sociales (ver la Tabla 6.1 en el capítulo anterior). A primera vista, una potencial excepción es la respiración. Pero, una investigación produjo algunas pruebas que corroboran su inclusión: las enfermedades psicológicas pueden surgir a partir de la dificultad para respirar en el asma y la bronquitis. La Tabla 7.1 muestra cómo cada una de las bases de las emociones puede también dar lugar a la correspondiente enfermedad mental.

Tabla 7.1. Ejemplos de síntomas comunes y sus raíces en sensaciones corporales y emociones (véase la Tabla 6.1 en relación con la ontología subyacente)

• trastornos relacionados con la comida, la bebida y la respiración, p. ej., anorexia nerviosa

• trastornos referidos a la salud, p. ej., hipocondría

• la evitación de sustancias nocivas, p. ej., transtorno obsesivo-compulsivo

• el mantenimiento de la vida y el logro de los objetivos, p. ej., hipomanía

• la evitación de situaciones peligrosas, p. ej., fobia

• el mantenimiento de las relaciones entre los cuidadores y los niños, p. ej., depresión

• relaciones sexuales, p. ej., impotencia psicológica

• relaciones con los iguales, p. ej., ansiedad social

Cuando sentimos una emoción intensa respecto a un suceso, nuestra mente se concentra en ella. Razonamos sobre el suceso para intentar dar sentido a la emoción, y decidir qué hacer respecto a él, si es que hacemos algo. El efecto de este razonamiento consciente, en el que se desarrollará mucha práctica, puede ser mantener nuestra enfermedad mental. Nos centra en la situación que conduce a la emoción anómala. Nos capacita para elaborar nuestra evaluación cognitiva inicial, y así puede ampliar las circunstancias que desatan la transición inconsciente. La práctica que adquirimos haciendo estas inferencias predice que nuestro razonamiento sobre temas pertinentes para nuestra emoción superaría tanto al razonamiento de los demás sobre esos mismos temas, como a nuestro propio razonamiento sobre otros temas. Esta predicción es crucial debido a que diferencia esta teoría de la opinión según la cual el razonamiento defectuoso es la causa de la enfermedad mental. Vuelvo ahora a algunos ejemplos de enfermedad mental para ilustrar la teoría.

* * *

Quien tiene un transtorno obsesivo-compulsivo a menudo pasa por la vida obligado a llevar a cabo conductas rituales de limpieza. Una paciente, conocida como “Ruth”, dedicaba cada día horas a lavarse. Después de utilizar el inodoro se restregaba cada dedo, luego los dos lados de las manos y luego los brazos. Llevaba a cabo esta tarea primero con jabón y luego con desinfectante. Luego limpiaba la taza y el lavabo. Y después repetía el ritual del lavado de manos. Si todavía se sentía contaminada, se daba una ducha. Se lavaba las manos después de cocinar y de limpiar la casa, pero solo en la medida en que estos trabajos le hacían sentirse sucia. Se lavaba las manos tres o cuatro veces cada hora, y se duchaba seis o siete veces al día. Sin embargo, no toda suciedad era “mala”. Ruth disfrutaba pintando cuadros, y no ponía pegas si tenía manchas de pintura. Eran suciedad “buena”. Ruth no tenía vida social, y sentía ansiedad en las interacciones sociales. Temía que pudiera perder el control de sí misma. El origen de las compulsiones de Ruth se remontaba a su infancia. Llevaba a cabo sus rituales de lavado de manos desde los cinco años. Fue por vez primera a un

terapeuta a los diez años.

De acuerdo con la teoría de la hiperemoción, la causa inmediata de los rituales de limpieza es la evaluación de ciertos hechos –la suciedad doméstica, la defecación y similares– como repugnantes. Estos hechos provocan sensaciones de repugnancia en la mayoría de las personas. En Ruth, sin embargo, provocan una transición inconsciente a una hiperansiedad respecto al peligro de contaminación. Esta inferencia corresponde a causación por contacto físico, y es un ejemplo de magia contagiosa. Ruth sería negligente si no tomara precauciones contra la contaminación. Y, como mucha gente con este problema, infiere de manera consciente que hace mal poniéndose en riesgo de contaminarse, y por tanto se siente culpable. Induce sensatamente que el peligro no es insuperable e incluso razona que no existe un mecanismo para el daño. Pero estas inferencias aumentan sus ideas sobre la probabilidad de contaminación. Mientras exista el más leve riesgo, si no se lavara descuidaría sus obligaciones. Si el lavado no es completo, se mantiene el riesgo de contaminación. Su madre le había inculcado un régimen de hiperlimpieza y pulcritud durante su infancia. La adquisición de la transición inconsciente fue, supongo, una consecuencia de ese régimen.

Si sufrimos de compulsiones obsesivas, nuestras cavilaciones tienen un patrón característico que se ilustra con la escena del inicio del capítulo. Pero, como cualquier pensamiento, nuestro razonamiento no es determinista. Podemos omitir un paso o perseverar en una preocupación determinada durante un cierto tiempo. Nuestro razonamiento se inicia con un pensamiento sobre un peligro potencial, y parece no venir de ninguna parte, lo mismo que muchos pensamientos de la vida diaria. Contiene una pizca de ansiedad racional, pero conduce a una transición inconsciente a la ansiedad extrema. Luego, como resultado de esta emoción, nos centramos en el improbable peligro. Nuestra hiperansiedad puede llevarnos a sobrevalorar lo severo y lo probable que es el peligro. Siguiendo un patrón característico de razonamiento, buscamos contraejemplos para la peor situación posible y luego ejemplos a su favor. Esta forma dialéctica de razonamiento nos conduce normalmente a aceptar el peligro como real. Y así seguimos en un bucle indefinido de pensamientos y sensaciones, oscilando entre un ejemplo que produce el peor resultado y un ejemplo en su contra.

Algunos psicólogos han propuesto que los sujetos con desorden obsesivo-compulsivo razonan de una forma anormal en la que de algún modo fusionan

pensamientos con acciones o sucesos. En contraste, la presente teoría postula que los pacientes razonan normalmente, y que sus muchas inferencias sobre las fuentes de su ansiedad les convertirían en razonadores expertos en relación con ese tema. Como muestra el ejemplo sobre el SIDA, son capaces de construir amplias secuencias de inferencias. Asimismo, se basan en mecanismos normales para considerar posibilidades. Siempre que pensamos sobre cualquier asunto con detalle, es probable que recojamos más pruebas que aumentan su probabilidad. Las cavilaciones de los pacientes sobre ejemplos y contraejemplos de un peligro lo ampliarían con más detalles, y por tanto aumentarían su probabilidad subjetiva. A lo largo de una serie de episodios que, en ausencia de tratamiento, podrían durar toda una vida, el riesgo de contaminación se hace más plausible y más general. Esa es la razón por la que Ruth llegó a dedicar tantas horas del día y de la noche a su lavado ritual.

* * *

Quien sufre de hipocondría pasa la vida preocupado por estar enfermo, quizá mortalmente enfermo. Thomas Carlyle, autor en el siglo diecinueve de The French Revolution, fue un famoso hipocondríaco. Parecía entender algo sobre su enfermedad, e imagino que tuvo este diálogo con su médico:

Carlyle: Estoy preocupado porque puedo ser un hipocondríaco.

Doctor: No se preocupe, solo se lo está imaginando.

Un caso más típico de un episodio de hipocondría le ocurrió a un joven ejecutivo que estaba a punto de marcharse al extranjero con su familia para iniciar un nuevo trabajo. Poco antes de la partida, se despertó con visión borrosa, y tuvo un temor repentino y abrumador a haber desarrollado esclerosis múltiple. Estaba tan preocupado que fijó una cita urgente con su médico. El doctor le hizo un completo chequeo físico, incluido un examen oftalmoscópico del ojo. Después del escepticismo inicial del paciente, el doctor fue capaz de explicarle la relación

entre sus síntomas y sus preocupaciones respecto a pasar un tiempo en el extranjero. El paciente, más tranquilo, aceptó el diagnóstico del doctor. Muchos pacientes, sin embargo, se resisten a la idea de que tienen una enfermedad mental, y su hipocondría, como la de Carlyle, dura toda la vida.

La hipocondría empieza con ciertas sensaciones corporales. Estas sensaciones –dolores de estómago, ligeros mareos, náuseas, etc.– pueden interpretarse como signos de enfermedad, y esta evaluación inicia una transición inconsciente hacia un grado de ansiedad anómalo. Cualquiera puede experimentar una ligera hipocondría. Siente un dolor en el pecho, y piensa: quizá estoy teniendo un ataque al corazón. Pero, el que es fuerte tiende a pensar: puede ser solo una indigestión, voy a esperar y a ver. El dolor desaparece, y retoma la vida normal. El entorno puede contribuir a la aparición de la hipocondría. Los estudiantes de medicina son famosos como grupo que con frecuencia tiene síntomas de hipocondría. De la misma forma, al consultar una enciclopedia médica, la descripción muy gráfica de una enfermedad puede provocar la inferencia de que nosotros también la estamos sufriendo. Y hay quien se centra en las sensaciones corporales hasta tal punto que desarrolla ansiedad respecto a las enfermedades mortales. Cavilamos sobre esa posibilidad y nuestras sensaciones corporales nos mantienen centrados en la peor de las situaciones posibles, y así, como muestra la escena al inicio del capítulo, no oscilamos entre un argumento en favor del peor resultado y otro en su contra. En lugar de eso hacemos inferencias para justificar nuestros temores, y estas inferencias exacerban nuestros problemas.

Una consecuencia de la hipocondría es que nos centremos tanto en las sensaciones corporales que pequeñas sensaciones nos produzcan ansiedad. Otra consecuencia es que tengamos encuentros frecuentes con los médicos –hasta el punto de convertirnos en expertos en la enfermedad que creemos tener–. Podemos pasar el día entero pensando en ella. Nuestra incertidumbre es similar a la apuesta de Pascal sobre la existencia de Dios –creer en Dios es racional, sostenía, porque no pasa nada trascendente si la creencia es errónea, pero si el escepticismo es erróneo tiene consecuencias desastrosas–. De manera similar, creemos que estamos enfermos. Por consiguiente, para nosotros es racional consultar a un médico porque no pasa nada trascendente si nuestra creencia es errónea, pero si el escepticismo respecto a la enfermedad es erróneo, el no consultar a un médico es desastroso. Vamos al médico, que no nos encuentra ningún problema. Seguimos teniendo nuestros síntomas físicos, y por tanto podemos rechazar el diagnóstico del doctor.

* * *

La teoría de la hiperemoción postula que las enfermedades psicológicas son causadas por transiciones inconscientes que dan lugar a emociones básicas anómalas, y se mantienen mediante patrones de razonamiento consciente característicos. Ya he presentado pruebas que demuestran el razonamiento inconsciente, y por tanto la pregunta crucial es si las enfermedades psicológicas son el resultado de transiciones a emociones básicas. Es difícil, pero no imposible, obtener datos respecto a este punto.

Anteriormente, revisé los hallazgos sobre la mediación de la amígdala en las emociones básicas. De la teoría se deduce que la amígdala debería activarse siempre que los hechos hagan surgir los signos y síntomas de una enfermedad psicológica. Esta predicción ha sido confirmada en diversos estudios sobre imágenes cerebrales, utilizando la resonancia magnética (el procedimiento RMf que describí en el Capítulo 2). Han mostrado un alto grado de actividad en la amígdala cuando se presentan estímulos relevantes a sujetos que sufren de transtorno obsesivo-compulsivo, hipocondría, fobia y transtorno por estrés postraumático. De la misma forma, la actividad en la amígdala correlaciona con depresiones de varios tipos y con su gravedad.

Un estudio epidemiológico a pequeña escala corrobora también el papel de las emociones básicas en la aparición de enfermedades psicológicas. Veinticuatro psiquiatras que trabajaban en Roma y Verona completaron un cuestionario de 15 preguntas respecto a los pacientes que habían diagnosticado más recientemente como obsesivos, agorafóbicos, hipocondríacos o deprimidos. El cuestionario incluía preguntas sobre el diagnóstico, el sexo y otros detalles del paciente, y sobre si el paciente tenía algunos otros síntomas de enfermedad. También preguntaba si los pacientes recordaban el inicio de su enfermedad, y si era así, qué emoción habían sentido entonces. Esta cuestión incluía una lista de control con cinco términos que se referían a emociones básicas (ansiedad, repugnancia, temor, enojo y tristeza) y cinco términos que se referían a emociones complejas (culpa, vergüenza, orgullo, bochorno, envidia). Las palabras estaban en un orden aleatorio, y los psiquiatras podían añadir otros términos referidos a emociones, si era necesario. Ninguno conocía la intención del estudio, la teoría de la

hiperemoción, o, ni siquiera, que esas eran las preguntas críticas. Todos los psiquiatras rellenaron el cuestionario, y proporcionaron datos de cuarenta y nueve hombres y cincuenta y siete mujeres. Entre los pacientes que recordaban la aparición de sus enfermedades (todos menos cinco), la mayoría informaba de la existencia en su inicio de una emoción básica (más del ochenta por ciento), y la única excepción era la experimentación de culpa en casos de transtorno obsesivo-compulsivo o depresión (doce por ciento de los casos). Puede que la culpa haya aparecido posteriormente en la enfermedad, y los recuerdos de los pacientes pueden haberles engañado. Incluso así, los resultados corroboraban la predicción de que las emociones básicas son, con mucho, el elemento más frecuente en el inicio de una enfermedad mental.

* * *

Quienes tienen una tendencia hacia una enfermedad mental deberían centrarse en temas pertinentes para su enfermedad y, como resultado, razonar sobre ellos con más precisión que los participantes de control. Sin embargo, con materiales neutrales, la diferencia debería desaparecer. No hubiera sido ético someter a los pacientes a un régimen que probablemente les provocara sus signos o síntomas, y por ello, en lugar de hacer eso, evaluamos a estudiantes universitarios con tendencias hacia las enfermedades psicológicas, que ni habían pedido tratamiento ni habían sido diagnosticados como enfermos. La predicción crucial era que esos participantes deberían razonar mejor que los participantes de control, pero solo con materiales relacionados con sus potenciales enfermedades; con contenidos neutrales, la diferencia entre los dos grupos debería desaparecer. Esta predicción es contraria a aquellas descripciones de las enfermedades psicológicas que sitúan sus causas en el razonamiento defectuoso.

Llevamos a cabo dos experimentos para evaluar la predicción. Nuestro sondeo epidemiológico sugería que dos comparaciones reveladoras deberían ser, en primer lugar, la realizada entre aquellos con tendencia hacia las compulsiones obsesivas y un grupo de control, y en segundo lugar, la realizada entre aquellos con tendencia hacia la depresión y un grupo de control. Dado que los participantes con tendencias hacia las obsesiones o hacia la depresión también era probable que fueran propensos a la ansiedad, los dos experimentos

examinaron el razonamiento sobre la culpa en aquellos participantes con tendencia hacia el transtorno obsesivo-compulsivo, y el razonamiento sobre la tristeza en aquellos participantes con tendencia a la depresión.

En cada prueba, la tarea de los participantes era leer una escena corta que conducía a una proposición concreta, y luego enumerar lo que era posible y lo que era imposible de acuerdo con esa proposición. Dada una oración como por ejemplo:

Suena la alarma o estoy cansado, o ambas cosas.

se decía a los participantes que imaginaran que la oración era verdadera, y enumeraran lo que era posible en función de las posibles combinaciones o de si la alarma sonaba o no y de si se sentían cansados o no. La actuación correcta sería enumerar las siguientes posibilidades completas, que dejan claro el estatus de ambas cláusulas de la disyunción en cada posibilidad:

Suena la alarma. No estoy cansado.

No suena la alarma. Estoy cansado.

Suena la alarma. Estoy cansado.

Los participantes tenían que enumerar luego de la misma forma lo que era imposible:

No suena la alarma. No estoy cansado.

En el primer estudio, se pasó un test de tendencias obsesivo-compulsivas a 290 estudiantes, y tomaron parte en el experimento el cinco por ciento que tenía las puntuaciones más altas y el cinco por ciento que tenía las puntuaciones más bajas. Estos dos grupos se subdividieron posteriormente en dos. En el primer caso, los participantes tenían que enumerar las posibilidades para proposiciones con contenidos diseñados para provocar una emoción de culpa. La historia que precedía a la proposición se refería a la culpabilidad del protagonista, p. ej.:

Supongamos que estoy en mi casa con unos amigos. Decidimos reunirnos con otros amigos en un bar. Dejamos la casa bromeando entre nosotros, pero olvido cerrar la ventana del baño.

La oración para la que había que enumerar las posibilidades era:

Suena la alarma antirrobo y me siento culpable.

En el segundo caso, los participantes tenían que enumerar posibilidades con oraciones referentes a la depresión, p. ej.:

Suena la alarma antirrobo y me siento deprimido.

o con proposiciones neutrales. Los participantes de cada grupo llevaron a cabo la tarea cuatro veces en un orden aleatorio diferente: en dos de las historias la proposición de la prueba se basada en “y” y en dos de las historias la proposición de la prueba se basada en “o”.

Los participantes obsesivos enumeraban como posibilidades verdaderas para la oración sobre la culpa más posibilidades completas correctas (sesenta y tres por ciento) que los participantes que no eran obsesivos (veintitrés por ciento). Pero, cuando la historia y la proposición tenían, o bien un contenido relacionado con la depresión, o bien uno neutral, no había diferencia fiable en la enumeración correcta de las posibilidades completas entre el grupo de los obsesivos (siete por ciento) y el de los no obsesivos (diecinueve por ciento). Al enumerar lo que era imposible se producía un patrón similar. Cuando examinamos las posibilidades que los participantes describían de forma incompleta ambos patrones se mantenían. Una predisposición al transtorno obsesivo-compulsivo puede mejorar el razonamiento sobre contenidos relevantes respecto al transtorno.

Un segundo estudio mostró que los participantes con una tendencia a la depresión, mostrada por un cuestionario de evaluación de depresión, también manifestaban un aumento comparable en su habilidad para razonar sobre materiales que se referían a asuntos depresivos. Es decir, razonaban más exactamente que los participantes de control sin tendencias depresivas. Sin embargo, la diferencia desaparecía con materiales neutrales o con materiales referidos a la culpa. La moraleja de los estudios es que el razonamiento de sujetos con propensión hacia una enfermedad mental en relación con asuntos pertinentes a su enfermedad, es mejor que el razonamiento de los demás o que su propio razonamiento sobre otros asuntos.

* * *

Quien sufre una enfermedad mental hace transiciones inconscientes a emociones básicas. Puede saber la causa de sus emociones, pero no conoce las razones de su anómala intensidad. Cuando entra en un ascensor, a un electricista de Manhattan que hizo algunos trabajos para mí le invade la ansiedad. ¿Por qué? Dice que teme que perderá el control de sí mismo si las puertas se cierran y el ascensor empieza a moverse. ¿Por qué? No dispone de ninguna explicación. En el pasado, este tipo de vacío se llenaba con diversas teorías. Los psicoanalistas lo llenaban con apelaciones a deseos reprimidos alejados del objeto inmediato de la emoción. Los terapeutas cognitivos dudan del papel de estos factores inconscientes. Hacen hincapié en que un interrogatorio más cuidadoso de los

pacientes da lugar a pensamientos que revelan que la fuente de las enfermedades está en inferencias defectuosas. Pero, de acuerdo con la presente teoría, los comentarios de los pacientes no aclaran la causa de la hiperintensidad de sus emociones, sino simplemente sus pensamientos sobre el objeto o situación que precipita el que surja la emoción. Su razonamiento no es la causa de su enfermedad, sino su consecuencia. Ciertamente, quienes tienen una propensión a la enfermedad mental razonan mejor que los participantes de control, aunque solo en asuntos que se refieren a la enfermedad. Desgraciadamente, sus conclusiones tienden a exacerbar su condición en lugar de a remediarla. El patrón de razonamiento depende de la emoción – sea miedo a la contaminación, una enfermedad mortal o un ataque de pánico–. Sus resultados refuerzan el que se centren en la situación que la provoca, y pueden conducir a una generalización de los síntomas.

¿Y la moraleja terapéutica? Si tenemos una enfermedad mental, el sistema de transiciones a las emociones básicas necesita ser reeducado. Entre las muchas formas de terapia existen ya varios procedimientos que pueden ayudar a reeducar el sistema. Si se trata de reducir una transición inconsciente hacia la ansiedad, la terapia de “exposición” parece ser la más efectiva. Después de una preparación inicial, se expone a los pacientes a sucesos que provocan su emoción anómala. A lo largo de todo el tratamiento, tienen que abstenerse de las acciones que llevan a cabo para reducir su ansiedad –evitación o escape en caso de agorafobia o transtorno de pánico, acciones compulsivas en caso de desorden obsesivo- compulsivo y autoafirmación en el caso de la hipocondria–. Su ansiedad se calma en una hora, y es menos intensa en su siguiente encuentro con la situación relevante. La terapia funciona. Una vez que nos hemos enfrentado a nuestros peores miedos y los hemos controlado, es probable que revisemos los modelos que tenemos de nosotros mismos y que recuperemos nuestra anterior confianza.

Notas al pie

[5] Nota de la trad.: El autor hace un juego de palabras entre bugs in the program (virus o bichos en el programa) y la expresión bats in the belfry (murciélagos en el campanario), que es el equivalente a la expresión en castellano “estar como

una cabra”.

Parte III

Cómo hacemos deducciones

¿No se podría decir incluso que la
serpiente fría del árbol venenoso
era l'esprit de géometrie,
y que Eva y Adán hasta la Caída
fueron totalmente ilógicos
pero al probar el fruto
echó raíces el pecado silogístico?

Auden

CAPÍTULO 8

Únicamente conexiones

En la novela policíaca de Agatha Christie Asesinato en el Oriente Express, el señor Ratchett (alias Casetti, con fama de secuestrador) muere apuñalado en el tren que va de Estambul a Calais. Hercules Poirot, una inteligencia sobrehumana como no ha habido otra, infiere que en el asesinato ha intervenido la mano de más de una persona, y tiene que averiguar quién, entre los trece sospechosos del vagón Estambul-Calais, apuñaló a Ratchett. La Tabla 8.1 presenta una forma de describir la situación, si bien es cierto que no es la de Poirot, quien se apoya en una cierta dosis de conjeturas. Dado que hay trece sospechosos, que en cada caso pueden o no haber tomado parte en el asesinato, hay más de ocho mil posibilidades distintas en las que más de una persona apuñaló a Ratchett. Podrá usted entender ahora por qué el razonamiento es inabarcable: si hubiera cien sospechosos el número de posibilidades sería astronómico. Un lógico puede establecer las posibilidades respecto al apuñalamiento de Ratchett en una inmensa tabla de verdad, y luego utilizar el significado de cada premisa para eliminar posibilidades de la tabla. La primera premisa, por ejemplo, descarta todas las posibilidades en las que la Condesa colaboró en el asesinato pero Cyrus no. Cuando se han utilizado todas las premisas para eliminar posibilidades, lo que quede debe ser la solución. La tabla de verdad del lógico es imposible de utilizar sin la ayuda de lápiz y papel, y deberíamos ser bastante ingeniosos para que se nos ocurriera ese método.

En la Parte I de este libro yo sostenía que captamos los significados de las proposiciones y luego utilizamos esos significados y nuestro conocimiento para construir modelos mentales sobre la significación de las proposiciones, es decir, las posibilidades a las que se refieren. Lo que voy a describir en esta parte del libro es cómo hacemos deducciones partiendo de premisas conscientes para llegar a conclusiones conscientes –deducciones conscientes, para abreviar–. Como sostuve en la Parte II, todos los procesos de pensamiento son inconscientes, y por tanto estas deducciones conscientes dependen de procesos inconscientes. La inferencia sobre el apuñalamiento de Ratchett depende solo de

las conexiones que las premisas establecen entre varias proposiciones simples, tales como, por ejemplo: colaboró la Condesa y colaboró Cyrus. No tenemos necesidad de hurgar en el interior de estas proposiciones para inferir quién tomó parte en el asesinato. Desde un punto de vista lógico, las inferencias de este tipo son las más simples que se pueden hacer, y la rama de la lógica que es más relevante para ellas se conoce como cálculo “proposicional” o “sentencial”. También hacemos deducciones que giran en torno a lo que hay en el interior de las proposiciones. Estas deducciones son más complicadas lógicamente, y a ellas se refieren los capítulos posteriores; pero este capítulo va a describir cómo hacemos deducciones que dependen solo de conexiones lógicas entre proposiciones. Es un enigma que los psicólogos han estado investigando durante cuarenta años, y este capítulo presenta la solución basada en modelos mentales y la solución a quién asesinó al señor Ratchett. Entenderá usted mejor estas soluciones si empiezo con un poco de lógica ligera.

Tabla 8.1. ¿Quién apuñaló a Ratchett en el Oriente Express?

Ratchett fue apuñalado doce veces, y esta es una descripción de quién colaboró en apuñalarle:

Si colaboró la Condesa entonces Cyrus también colaboró.

Colaboró Edward o colaboró la Condesa.

Si colaboró Pierre si lo hizo Hubbard entonces Greta también colaboró.

Greta no colaboró o lo hizo Hubbard.

Si no colaboraron tanto Hildegarde como Hector entonces no colaboró Hubbard.

Colaboró Edward o colaboraron la Princesa, Mary, y el Coronel.

Si colaboraron Hector o Mary entonces, o Pierre colaboró o el Coronel no lo hizo.

Colaboró la condesa o bien colaboraron tanto Hildegarde como Héctor.

Si la Condesa no colaboró entonces colaboraron Pierre, Greta y Edward.

Si el Conde no colaboró entonces Hubbard y Héctor, a la vez, no colaboraron

Colaboraron Cyrus y la Princesa o bien colaboraron el Conde y la Condesa.

Si no colaboraron tanto Mary como el Coronel entonces colaboró la Condesa.

Estas premisas muestran que todos menos uno tomaron parte en el apuñalamiento. Por tanto, ¿quién lo hizo o más bien quién no lo hizo? Encontrará la respuesta al final del capítulo.

▪

* * *

Puede que recuerde del Capítulo 1 que una oración como por ejemplo: “Ahora yo puedo oírte”, tiene un significado que entendemos a partir de nuestro conocimiento del idioma, pero su significación –la proposición concreta a que se refiere– depende de quién es el que habla, quién es el que escucha y el momento de la emisión. Sin embargo, la lógica considera que las oraciones expresan proposiciones únicas sin tener en cuenta el contexto. Y por eso voy a considerar oraciones que tienen una significación cercana a su significado, es decir, hasta donde es posible son independientes del contexto. Podemos ahora preguntarnos qué significa decir que una conectiva “sí” tiene un significado lógico.

Un paso importante para responder a la cuestión fue dado por el lógico del siglo diecinueve George Boole. Inventó el “álgebra booleana”, que es el álgebra del y, o y no, y aparece en todas partes, desde el diseño de ordenadores al cálculo proposicional. (Mary Boole, su mujer, escribió que el problema de Macbeth era que no dominaba el álgebra booleana.) Las conectivas lógicas son aquellas palabras cuyos significados lógicos se pueden descomponer en y, o y no. Estamos llegando al final, excepto porque ahora tengo que definir los significados lógicos de y, o y no.

Para volver desde Boole a Estambul:

Poirot se metió el billete en el bolsillo y subió al Oriente Express.

Entendemos que los dos sucesos se producen en un orden temporal: primero Poirot se metió el billete en el bolsillo, y luego entró en el tren. Sabemos que hay un orden temporal, porque una oración con las dos cláusulas en orden opuesto dice algo diferente:

Poirot subió al Oriente Express y se metió el billete en el bolsillo.

Una buena evaluación del significado lógico de y es que el orden de las dos cláusulas que conecta no tiene efecto sobre la significación de la oración, p. ej.:

Poirot tenía un billete y viajaba en el Oriente Express.

Esta oración es cierta con tal de que sean ciertas las dos proposiciones que expresa; y su valor de verdad, si es verdadera o falsa, no depende de nada más. De ahí que, suponiendo que es cierto que Poirot tenía un billete y que es cierto que viajaba en el Oriente Express, la conjunción de las dos proposiciones también sea cierta. Este es el núcleo de lo que significa una conectiva lógica: el valor de verdad de la proposición basada en la conectiva depende solo del significado de la conectiva y de los valores de verdad de las proposiciones individuales que conecta. Los lógicos tienen en su jerga un término para expresar esta idea. Dicen que la conectiva tiene “función de verdad”. Una conectiva lógica tiene “función de verdad”.

Ahora hemos llegado al final. Una conectiva lógica entre proposiciones da lugar a una proposición cuya verdad depende de la verdad de las proposiciones que la constituyen. La naturaleza de la dependencia varía de una conectiva a otra. El significado lógico de y es que la conjunción que forma es válida suponiendo que todas las proposiciones de la conjunción son verdaderas; si alguna de ellas es falsa, entonces la conjunción también es falsa. El significado lógico de o es que la disyunción que forma es verdadera suponiendo que al menos una de las proposiciones de la disyunción es verdadera; si todas son falsas, entonces la disyunción también es falsa. ¿Qué pasa con no? No es una conectiva, pero su significado lógico es sencillo: una oración negativa es cierta suponiendo que la proposición que se niega es falsa, y viceversa. Por ejemplo, la proposición Poirot no tiene un billete es verdadera, si la proposición Poirot tiene un billete es falsa. Cuando un personaje de una de las obras de Agatha Christie dice, “No podemos no llamar a la Policía”, sus dobles negaciones son lógicas, y se anulan para producir una afirmación: debemos llamar a la policía. El filósofo de Oxford J. L. Austin comentó una vez en una clase que, a diferencia de las negaciones, las dobles afirmaciones no se anulan para producir una negación, “Sí, sí”, dijo otro

filósofo, el difunto Sidney Morgenbesser, desde el fondo de la sala de clase.

No necesitamos tratar las tres conectivas de Boole como primitivas. Por ejemplo, podemos definir cualquier conectiva en relación con las dos conectivas, no e y. Ciertamente, con ingenio podemos definir cualquier cosa con una única conectiva, ny, que es una forma abreviada de ambos ___ y ___ no. También podemos definir nuevas conectivas a partir de nuestra serie de tres. Puede que usted se haya dado cuenta, por ejemplo, de que la definición que di más arriba para “o” correspondía a su significado incluyente, el cual permite que la disyunción sea verdadera si cada proposición de la disyunción es verdadera. Podemos usar las tres conectivas para definir el significado excluyente de “o”. Una disyunción excluyente, como por ejemplo:

O colaboró Edward o colaboró la Condesa, pero los dos no.

es equivalente a la proposición:

Colaboró Edward o colaboró la Condesa, y no ambas cosas: colaboró Edward y colaboró la Condesa.

Por consiguiente, la disyunción excluyente también tiene un significado lógico.

De forma similar podemos definir un significado lógico para “si”. La oración:

Si no colaboró la Condesa entonces lo hizo Pierre.

significa:

Colaboró la Condesa o lo hizo Pierre, o ambos.

En su significado lógico el condicional es compatible con tres posibilidades: la Condesa no colaboró y Pierre sí, la Condesa colaboró y Pierre no, la Condesa colaboró y Pierre colaboró. La única posibilidad que excluye el condicional es que ni la Condesa ni Pierre colaboraran. Las tres posibilidades que permite el condicional son las mismas que las de la disyunción incluyente. Esta equivalencia, como veremos, no se mantiene para todos los condicionales, sino solo para aquellos con un significado lógico.

Los lógicos pueden utilizar tablas de verdad para determinar si una inferencia es válida o no: es válida si su conclusión debe ser verdadera en el caso de que las premisas sean verdaderas. Pero también pueden definir una lógica con reglas formales de inferencia que les permite probar que la conclusión se sigue de las premisas. Como ya mencioné, uno de los méritos de la lógica del siglo veinte es el descubrimiento de que hay lógicas en las que no todas las inferencias válidas se pueden probar utilizando un conjunto consistente de reglas formales de inferencia. Sin embargo, la lógica de las tres conectivas –el cálculo proposicional– tiene la feliz propiedad de que todas las inferencias que son válidas sobre la base de los significados se pueden probar también en un sistema de reglas formales de inferencia consistentes, y viceversa.

* * *

La búsqueda de una prueba para el cálculo proposicional mediante el uso de reglas formales de inferencia puede llevar muchísimo tiempo, debido a que el campo es inabarcable. El uso de tablas de verdad también es problemático debido a que el tamaño de las tablas de verdad se duplica con cada proposición básica de la inferencia. Por eso es por lo que hay más de ocho mil posibilidades en la tabla de verdad para el asesinato en el Oriente Express. El primer paso para llegar a una teoría psicológica del razonamiento proposicional es considerar una versión más eficiente de las tablas de verdad, que he implementado en un programa de ordenador.

He mencionado ya varias veces los programas de ordenador, pero antes de describir este programa en concreto, es el momento de decir algo sobre el papel de los programas en la psicología cognitiva. Cuando los psicólogos teorizan sobre los procesos, un peligro es que den demasiadas cosas por supuestas, y por tanto sus descripciones sean imprecisas o difíciles de entender. Otro peligro es que invoquen a un homúnculos para proporcionar la explicación. Se arriesgan a situar un hombrecillo dentro del cerebro que hace todo el trabajo, y así mágicamente una rueda dentada de la máquina pasa a tener todos los poderes de la máquina en sí misma. Una defensa frente a la imprecisión y los homúnculos es implementar los componentes clave de una teoría en un programa de ordenador que funcione. Como escribió Lord Kelvin, el físico del siglo diecinueve:

Nunca me quedo satisfecho hasta que puedo hacer un modelo mecánico de algo. Si puedo hacer un modelo mecánico puedo entenderlo. Mientras no puedo hacer un modelo mecánico de todo el proceso no puedo entender...

Un programa de ordenador es un modelo mecánico. Sin embargo, incorpora demasiadas pequeñas suposiciones como para captarlas con los trazos gruesos de un experimento. Podríamos hacer otras pequeñas suposiciones distintas y llegar a un programa diferente que siga simulando la misma teoría. Por consiguiente, los experimentos evalúan las teorías, no los programas.

La programación de ordenadores no es un deporte para espectadores. Leer una descripción de un programa de ordenador es como leer la descripción de una muestra de punto de media. Nada es más impenetrable y menos divertido –a menos que uno quiera tejer el jersey real, o escribir el programa real–. Por tanto, puede que usted quiera avanzar rápidamente, captando solamente la idea general de cómo funciona el programa.

La entrada [input] del programa es un conjunto de premisas, como las del asesinato de Ratchett, y su salida [output] es el conjunto de posibilidades compatibles con ellas, seguida por una descripción parsimoniosa de las posibilidades. En el caso del apuñalamiento de Ratchett, el programa revela la identidad de la única persona que no participó en el asesinato. Representa el

conjunto de posibilidades de forma completa y totalmente explícita. Construye estos modelos completos de las posibilidades respecto a la primera premisa, y luego los actualiza formando una conjunción lógica de cada uno de ellos con cada uno de los modelos completos de la segunda premisa, y así sucesivamente, hasta que ha tenido en cuenta todas las premisas. El mecanismo para la conjunción es simple. Si dos modelos contienen proposiciones contradictorias, el resultado es cero, que es el modelo nulo o vacío que el programa utiliza para representar autocontradicciones. Si dos modelos están libres de contradicción, el resultado es un modelo que representa todas las proposiciones contenidas en ambos modelos. Y, si uno de los modelos en una conjunción es el modelo nulo, el resultado es el modelo nulo. La Tabla 8.2 describe estas reglas y las ilustra en relación con las dos primeras premisas en el problema sobre el asesinato de Ratchett. El resultado son cuatro modelos, lo que es menos oneroso que la tabla de verdad de las premisas. En general, el programa representa solo modelos de las posibilidades consistentes con las premisas y por tanto es más eficiente que el uso de las tablas de verdad, que también representan las posibilidades que las premisas descartan. Resuelve el misterio del asesinato sin tener que construir en ningún caso más de setenta y cinco posibilidades al mismo tiempo frente a la tabla de verdad que representa más de ocho mil posibilidades.

Tabla 8.2. El mecanismo para hacer una conjunción de modelos completos

Las premisas:

Si colaboró la Condesa entonces Cyrus también colaboró.

Colaboró Edward o colaboró la Condesa.

tienen respectivamente los siguientes modelos completos:

La Condesa colaboró Cyrus colaboró

La Condesa no colaboró Cyrus colaboró

La Condesa no colaboró Cyrus no colaboró

y:

Edward colaboró La Condesa no colaboró

Edward no colaboró La Condesa colaboró

Edward colaboró La Condesa colaboró

El mecanismo actualiza los modelos de la primera premisa formando una conjunción lógica de cada uno de ellos con cada uno de los modelos de la segunda premisa. En general hay tres procedimientos para formar conjunciones:

1. La conjunción de un par de modelos que contienen respectivamente una proposición y su negación da lugar a cero, el modelo vacío que representa las contradicciones, p. ej.:

La Condesa colaboró Cyrus colaboró

y:

Edward colaboró La Condesa no colaboró

dan lugar a cero.

2. La conjunción de un par de modelos que no contienen contradicción da lugar a un modelo que representa todas las proposiciones contenidas en ambos modelos, p. ej.:

La Condesa colaboró Cyrus colaboró

y:

Edward colaboró La Condesa colaboró

da lugar a la Condesa colaboró Cyrus colaboró Edward colaboró.

3. La conjunción de un par de modelos en los que al menos uno de ellos es el modelo nulo da lugar al modelo nulo, p. ej.:

La Condesa colaboró

y

cero

da lugar a cero

Por tanto, el resultado de conjugar los modelos de la primera premisa presentada anteriormente con los de la segunda premisa es:

La Condesa colaboró	Cyrus colaboró	Edward no colaboró
La Condesa no colaboró	Cyrus no colaboró	Edward colaboró
La Condesa colaboró	Cyrus colaboró	Edward colaboró
La Condesa no colaboró	Cyrus colaboró	Edward colaboró

Lo mismo que nosotros, el programa puede evaluar conclusiones dadas o extraer conclusiones por sí mismo. Extrae conclusiones parsimoniosas, es decir, aquellas que utilizan el menor número de conectivas. Los lógicos y los ingenieros han desarrollado diversos procedimientos para este problema, que se conoce como problema de ”minimización booleana”. Es importante porque los componentes de muchos circuitos electrónicos son booleanos. Por consiguiente, cuantos menos componentes tenga un circuito, más barata es su fabricación. Yo he

implementado un procedimiento que funciona descomponiendo conjuntos de modelos en subconjuntos, hasta poder construir una descripción apropiada. Se garantiza la obtención de una descripción lo más reducida posible. Sin embargo, el problema también es inabarcable. No importa lo eficiente que sea un procedimiento, más pronto o más tarde el número de modelos se hace demasiado grande como para computar una solución en un tiempo razonable.

* * *

Nuestro sistema de construcción de modelos construye los modelos de uno en uno, y, en la medida de lo posible, trata de no construir más que un modelo. La estructura y contenido de un modelo capturan lo que es común a las diferentes formas en que se puede producir la posibilidad. Cuando nos vemos obligados a intentar tener presentes varias posibilidades, la tarea es difícil, porque nuestra memoria de trabajo tiene una capacidad limitada. Una inteligencia sobrehumana no tendría esta limitación. Su memoria no sería un cuello de botella, y por tanto podría resolver el asesinato en el Oriente Express sin necesidad de lápiz y papel utilizando el mismo mecanismo que el programa de ordenador. No nos damos cuenta de nuestras limitaciones porque nuestro mundo social no es más complicado que nuestra habilidad para pensar en él, y nuestro razonamiento sobre el mundo físico es lo suficientemente bueno como para que sobrevivamos y nos reproduzcamos.

En general, una conclusión es válida si se cumple en todos los modelos de las premisas. Una conclusión no es válida si al menos uno de los modelos de las premisas es un contraejemplo respecto a ella. Como dije al comienzo del libro, la base de la racionalidad humana parece residir solo en este principio: una inferencia es válida si su conclusión se cumple en todas las posibilidades compatibles con las premisas, y no es válida si existe una posibilidad compatible con las premisas pero no con la conclusión. Las personas que razonan parecen captar este principio y darse cuenta de que una inferencia es falible si pueden pensar en un contraejemplo.

Una asunción fundamental de la teoría de modelos es el llamado principio de “verdad”: a diferencia de los modelos completos, que describí anteriormente, los

modelos mentales representan una proposición sencilla de las premisas solo cuando la proposición es verdad en esa posibilidad. La mente parece trabajar de esta forma con el fin de reducir la carga de la memoria de trabajo. Por consiguiente, a menos que suceda algo excepcional que lo invalide, se aplica el principio de verdad. Parece bastante inocuo, pero cuando lo haya explicado, llegaremos a su consecuencia más sorprendente e inesperada.

Como ejemplo del principio de verdad, consideremos la primera premisa del problema del apuñalamiento:

Si colaboró la Condesa entonces Cyrus también colaboró.

En su significado lógico, el condicional es compatible con las tres posibilidades que aparecen en la Tabla 8.2. Sin embargo, dado que, cuando razonamos a partir de un condicional, es difícil tener presentes tres posibilidades, nos centramos en la opción en la que son ciertas las dos proposiciones del condicional. Y así construimos el modelo mental de la posibilidad más destacada:

La Condesa colaboró Cyrus colaboró.

Si tuviéramos que construir solo este modelo, habríamos representado una conjunción: la Condesa colaboró y Cyrus colaboró. Pero nos damos cuenta de que no es necesario que la cláusula sí sea cierta: la Condesa puede no haber colaborado. En realidad aplazamos la construcción de un modelo explícito para esta posibilidad. Añadimos al modelo mental anterior uno que no tiene contenido explícito. Este modelo implícito actúa como un “contenedor” para recordarnos que hay otras posibilidades, y, al menos durante un tiempo, tenemos presente que en esas posibilidades es falso que la Condesa colaborara. Los modelos mentales del condicional según eso son:

La Condesa colaboró Cyrus colaboró

...

donde la elipsis indica el modelo implícito. Si recordamos lo que es falso en los modelos, podemos ampliar nuestros modelos mentales a los tres modelos completos de la Tabla 8.2.

La tarea de juzgar la verdad o falsedad es bastante diferente de la tarea de enumerar las posibilidades. Tenemos que relacionar los modelos de un condicional con los modelos del mundo para determinar si el condicional es verdadero o falso. Para llevar a cabo esta tarea nos apoyamos en los modelos mentales. Por consiguiente, tendemos a pensar que un condicional es verdadero solo en caso de que ambas cláusulas sean verdaderas, lo que corresponde al único modelo mental explícito del condicional. Sin embargo, para enumerar las posibilidades, solo necesitamos entender una oración, y así podemos ampliar nuestros modelos mentales a modelos completos.

La segunda premisa en el problema del apuñalamiento es una disyunción incluyente:

Colaboró Edward o colaboró la Condesa.

Sus modelos mentales de acuerdo con el principio de verdad son:

Edward colaboró

La Condesa colaboró

Edward colaboró La Condesa colaboró

En la primera posibilidad, la proposición de que la Condesa colaboró es falsa y por eso no está representada en el primer modelo mental. De manera similar, la proposición de que Edward colaboró es falsa en la segunda posibilidad y por eso no está representada en el segundo modelo mental. No debe usted confundir falsedad con negación: una oración negativa puede ser verdadera o falsa. Son las proposiciones falsas, sean negativas o afirmativas, las que no están representadas en los modelos mentales.

En estos diagramas sencillos de los modelos mentales, cada línea horizontal indica un modelo de una posibilidad independiente. Los modelos mentales reales, a diferencia de sus diagramas, no están formados por oraciones gramaticales, que utilizo por comodidad. Los modelos mentales representan las posibilidades de una forma tan icónica como sea posible, aunque no todo puede ser icónico. Como mostré anteriormente, ni la negación ni el estatus de un modelo como representación de una posibilidad pueden representarse en un icono. Ambas cosas tienen que representarse como símbolos. Nuestros modelos mentales son, por tanto, mucho más ricos que los diagramas anteriores.

* * *

Tenemos una tendencia natural a pensar en posibilidades. Por tanto, cuando extraemos una conclusión a partir de una serie de premisas, pensamos en las posibilidades compatibles con las premisas. Construimos los modelos mentales de una premisa, y luego los actualizamos de acuerdo con la siguiente premisa, y así sucesivamente. El programa de ordenador que he descrito también construye modelos mentales de las premisas así como modelos completos. Para hacer frente a la naturaleza incompleta de los modelos mentales es necesario un proceso adicional. Cuando una proposición como: Eduard colaboró, no aparece en un modelo de una premisa, pero aparece en otro modelo de la misma premisa, el programa interpreta su ausencia como su negación. Pero si tampoco aparece en ningún otro modelo de las premisas, su ausencia se considera consistente con su afirmación.

Como ejemplo ilustrativo, vamos a considerar de nuevo las dos primeras

premisas de nuestro problema del apuñalamiento:

Si colaboró la Condesa entonces Cyrus también colaboró.

Colaboró Edward o colaboró la Condesa.

Puede usted pensar por un momento qué posibilidades son compatibles con ellas. La mayoría de las personas piensan que son tres:

La Condesa colaboró	Cyrus colaboró	
		Edward colaboró
La Condesa colaboró	Cyrus colaboró	Edward colaboró

Lo que pasan por alto es la cuarta posibilidad mostrada en la Tabla 8.2 en la que solo colaboraron Cyrus y Edward. Quienes razonan basándose en modelos mentales no pensarán en esta posibilidad. Esta predicción es típica de la teoría de modelos.

Tabla 8.3. Modelos mentales y modelos completos para las proposiciones basadas en conectivas

Cada línea de los diagramas que aparecen a continuación representa una posibilidad compatible con una proposición.

Proposiciones	Modelos mentales	Modelos completos
A y B	A B	A B
Si A entonces B	A B	A B
	...	no-A B
		no-A no-B
A o B, o ambos	A	A no-B
	B	no-A B
	A B	A B
A o B, ambos no	A	A no-B

B no-A B

El contraste entre los modelos mentales y los modelos completos para las principales conectivas se resume en la Tabla 8.3. Estas diferencias y los principios para articularlas pueden llevar a diferencias radicales entre los modelos mentales y los modelos completos. Como resultado, es probable que cometamos errores al razonar sobre las conectivas porque dependemos de modelos mentales.

* * *

Dos inferencias opuestas aclaran la teoría. La primera inferencia es:

O bien Jane está arrodillada junto al fuego y está viendo la televisión o bien Mark está de pie junto a la ventana y está escudriñando el jardín.

Jane está arrodillada junto al fuego.

¿Se sigue que está viendo la televisión?

La mayoría de la gente dice: “sí”.

La segunda inferencia tiene la misma premisa inicial, pero en cambio va seguida por la proposición categórica:

Jane no está arrodillada junto al fuego.

y la pregunta es:

¿Se sigue que Mark está de pie junto a la ventana?

De nuevo, la mayoría de los participantes dicen: “sí”. Veamos qué predice la teoría.

La primera premisa de ambas inferencias es una disyunción excluyente de dos conjunciones. La teoría predice que deberíamos basarnos en modelos mentales. Por consiguiente, deberíamos construir un modelo que represente la primera conjunción, Jane está arrodillada junto al fuego y está viendo la televisión. Abrevio este modelo como sigue:

Jane: arrodillada viendo

Deberíamos construir un modelo análogo para la segunda conjunción:

Mark: de pie, escudriñando

Debemos combinar estos dos modelos de acuerdo con la disyunción excluyente de la premisa. Una disyunción excluyente tiene dos modelos mentales, que representan sus cláusulas en las posibilidades en las que son verdaderas. Por consiguiente, la disyunción excluyente tiene dos modelos:

Jane: arrodillada viendo

Mark: de pie escudriñando

En la primera inferencia, la premisa categórica tiene un único modelo mental:

Jane: arrodillada

Su conjunción con los modelos de la disyunción da lugar a:

Jane: arrodillada viendo

Y por ello los participantes deberían responder: sí, Jane está viendo la televisión. Este análisis puede sorprendernos por obvio.

En realidad, la inferencia es una falacia. El principio de verdad postula que representamos lo que es verdadero, no lo que es falso. Cuando escribí por primera vez un programa para simular la teoría, estudié su resultado y pensé que había un error en el programa. Busqué ese fallo durante medio día, antes de darme cuenta de que el programa era correcto, y que el error estaba en mi pensamiento. Lo que el programa revelaba era la discrepancia entre los modelos mentales, que no representan lo que es falso, y los modelos completos, que sí lo hacen. Por tanto, la teoría predice que deberíamos cometer errores sistemáticos para ciertas inferencias. Ciertamente, los errores resultan ser tan obligatorios en muchos casos que parecen ilusiones cognitivas, y por eso mis colegas y yo nos referimos a ellos como inferencias “ilusorias”.

La primera de las dos inferencias es ilusoria, y los modelos completos de las premisas revelan la conclusión correcta. Hay seis modelos completos de la disyunción excluyente, porque cuando una conjunción es verdadera, la otra conjunción es falsa, y hay tres formas en que una conjunción puede ser falsa. Puede ser falsa su primera proposición, puede ser falsa su segunda proposición y pueden ser falsas ambas proposiciones. De modo que, si Jane está arrodillada junto al fuego no se sigue que está viendo la televisión –la conjunción puede ser falsa siempre que la otra conjunción sea cierta, es decir, Mark está de pie junto a la ventana y escudriñando el jardín–. Por eso es por lo que la inferencia ilusoria es falsa.

El segundo de los problemas anteriores tiene la premisa categórica de que Jane no está arrodillada junto al fuego, y plantea la cuestión de si se sigue que Mark está de pie junto a la ventana. La mayoría de la gente responde, “sí”, que es una conclusión apoyada por los modelos mentales mostrados anteriormente. Esta inferencia es válida. Un experimento examinó una serie de inferencias ilusorias y problemas de control de este tipo. Los participantes acertaron con mucha más frecuencia en los problemas de control (setenta y ocho por ciento correctos) que en los problemas ilusorios (diez por ciento correctos), y en todos los participantes excepto uno mostraron esa diferencia. Confiar en la verdad es una forma sensata de enfrentarse a una memoria de trabajo limitada. En algunos casos, sin embargo, lleva a quien razona a la peligrosa ilusión de que capta posibilidades que en realidad no están a su alcance.

* * *

Las posibilidades están en el centro de la teoría de modelos, porque los modelos representan posibilidades. Sin embargo, el principio de verdad determina que los modelos mentales representan las posibilidades consistentes con las premisas, y dentro de cada posibilidad solamente representan las proposiciones sencillas de las premisas que son verdaderas. Sin embargo, los modelos completos también representan lo que es falso en las posibilidades. La ventaja de los modelos mentales es que contienen menos información. Ocupan menos espacio en la memoria de trabajo; pero pueden llevarnos por el mal camino. Somos propensos a las ilusiones porque no representamos el conjunto de posibilidades completas

consistentes con las premisas. Por ejemplo, pensamos en la verdad de cada proposición de una disyunción excluyente, pero al hacerlo dejamos de tener en cuenta la falsedad simultánea de la otra proposición. Es como si esa otra proposición hubiera dejado de existir. Las reglas sensibles a la forma lógica de las proposiciones no pueden predecir la aparición de las ilusiones. Estas reglas necesitan producir conclusiones válidas so pena de contradicción, y por tanto no tienen forma de predecir falacias sistemáticas.

Si las conectivas y, o y si tuvieran únicamente significados lógicos, el presente capítulo habría hecho una descripción completa de su interpretación. Pero, más adelante veremos cómo el conocimiento modula sus interpretaciones –hasta el punto de que dejan de tener significados lógicos–. Veremos también que la gente desarrolla una diversidad de estrategias para el razonamiento, que incluyen el uso de contraejemplos para refutar las inferencias no válidas. Finalmente, la solución al Asesinato en el Oriente Express es que todos excepto la Condesa apuñalaron al malvado señor Ratchett. El problema es trivial para el programa de ordenador cuando utiliza modelos completos.

CAPÍTULO 9

Soy mi propio abuelo: el razonamiento sobre identidades y otras relaciones

[1]

Ernest Harrison, miembro de un college de Cambrigde, participó como oficial naval en la Primera Guerra Mundial, y durante su etapa en el servicio activo se afeitó el bigote. Volvió de visita a Cambridge, y el director de su antiguo college, sin reconocerle, le preguntó durante la cena si era pariente de “nuestro querido Ernest Harrison”. Adoptando una cierta visión filosófica de las relaciones, él respondió: No.

La visión filosófica que sostuvo Harrison era que nadie puede ser su propio pariente. La lógica moderna es más sensata y admite que la identidad es una relación que se cumple entre cualquier objeto y ese mismo objeto. La lógica coincide con la vieja canción: “Yo soy mi propio abuelo.” Ciertamente, hacemos con frecuencia inferencias sobre relaciones, y la noción lógica de relación abarca no solo a los parientes, sino a cualquier cosa que vaya desde: más alto que hasta: la persona que emplea al marido de la estudiante de. Por ejemplo, podemos inferir la relación espacial entre el mueble de cajones y la mesa a partir de esta descripción del estudio del profesor Higgins (véase el Capítulo 2).

El mueble de cajones está a la derecha de la puerta.

La mesa está a la izquierda de la puerta.

No tenemos dificultad en llegar a la conclusión de que:

El mueble de cajones está a la derecha de la mesa.

La inferencia parece tan trivial como para no necesitar ninguna explicación. Pero, los psicólogos no están de acuerdo respecto a cómo la hacemos. Sin embargo, todos aceptan una cosa. Para extraer esta conclusión, en contraste con las inferencias descritas en el capítulo anterior, tenemos que examinar el interior de las proposiciones de las premisas con el fin de establecer una relación entre parte de una proposición y parte de la otra. Pueden existir relaciones entre entidades individuales, como en el ejemplo anterior, o entre conjuntos de entidades, como en la proposición, “No todos los muebles de cajones están a la derecha de todas las mesas”. Como puede imaginar, las relaciones entre entidades individuales son mucho más sencillas que las relaciones entre conjuntos, y por eso voy a dejar la descripción del razonamiento sobre conjuntos para más adelante.

* * *

¿Qué es una relación? Los lógicos han dado una respuesta precisa para esta pregunta, que yo voy a utilizar. Empecemos por la relación más alto que. Se establece entre un amplio número de parejas de entidades. Se establece, por ejemplo, entre esta pareja en el orden en que aparecen: Michael Jordan y George W. Bush, y también entre esta pareja en el orden en que aparecen: Monte Everest y Mont Blanc. Para utilizar un elemento de la terminología lógica, diré que la relación tiene dos “argumentos”, que son entidades como por ejemplo Jordan y Everest. La lógica define una relación como todos los pares de argumentos para los que se cumple la relación. En la vida cotidiana, el significado de la relación determina si existe o no para una pareja de argumentos determinada. El significado es un tema resbaladizo, no obstante, posteriormente en este mismo capítulo describiré cómo podemos analizar los significados de las relaciones espaciales. Un adjetivo como por ejemplo “feliz”, en “Pat es feliz” toma un argumento, como por ejemplo Pat, y establece una propiedad más que una relación. La lógica de las propiedades es sencilla, y fue descrita por primera vez

por Aristóteles. La lógica de las relaciones no es sencilla, y no fue totalmente definida hasta más de dos mil años después.

Los psicólogos siempre están tentados a reducir las relaciones a propiedades, porque las propiedades son más sencillas. El lenguaje nos ayuda a hacer esas reducciones. Por ejemplo, la proposición: “Dumbo es pequeño”, parece adscribir una propiedad a Dumbo. Pero, en realidad, afirma que Dumbo es pequeño en relación con el tamaño habitual de los elefantes. Por eso es por lo que no hay contradicción en afirmar: “Un elefante pequeño es grande.” Significa que un elefante que es pequeño para ser un elefante es grande para ser un animal típico. En la vida hacemos reducciones similares. Miramos al velocímetro del coche, y vemos que estamos viajando a sesenta millas por hora. La posición de la aguja en el dial muestra la velocidad instantánea del coche como si fuera una propiedad. Pero la velocidad de un coche depende de una relación entre distancia y tiempo.

Los psicólogos a menudo tratan las relaciones en la percepción como si fueran “características” –simplemente un sinónimo de propiedades–. Como versión general del conocimiento está condenada al fracaso, porque reducir las relaciones a propiedades tiene un coste. Para ver por qué, supongamos que queremos analizar la relación padre de como un conjunto de propiedades. Podríamos proponer este conjunto:

humano, adulto, varón, tiene-un-hijo.

La última de estas propiedades parece una relación disfrazada, pero no importa. Las relaciones se pueden combinar con otras relaciones para formar relaciones más complicadas. Esta, por ejemplo, es una relación de ese tipo:

la mujer del abuelo de.

Si la tratamos como un conjunto de propiedades no podemos distinguirla de la

muy distinta relación:

el abuelo de la mujer de.

Dorothy es la mujer del abuelo de Ben, pero no es el abuelo de nadie, y Ben tampoco es el abuelo de nadie. Las relaciones son relaciones: las reducimos bajo nuestra responsabilidad.

Un objeto puede entrar en muchas relaciones. Quizá, por ejemplo, usted está leyendo este libro, sentado en una silla, bebiendo una ginebra con tónica, comiendo foiegras, escuchando el sonido de las trompetas y esperando asombrarse. Todas estas relaciones se establecen entre solo dos entidades, o, para utilizar la terminología que introduje antes, dos argumentos: usted y algo más. Algunas relaciones se establecen entre tres argumentos, y en raras ocasiones entre cuatro, p. ej., Ben cambió su colección de sellos a Sara por una navaja. En teoría, se pueden establecer relaciones entre cualquier número de argumentos. Cuánto mayor sea el número de argumentos que tiene una relación más difícil de entender nos resulta, y el psicólogo Graeme Halford se refiere a este factor como “complejidad relacional”.

Las relaciones dan lugar a muchos tipos de inferencias, y los lógicos han identificado entre ellas varios modelos sistemáticos. Dicen que las relaciones son transitivas si dan lugar a este tipo de inferencia:

A está relacionado con B.

B está relacionado con C.

Por tanto, A está relacionado con C.

Otras relaciones son intransitivas. Por ejemplo, si Alan es el que está detrás de Bill en la fila, y Bill es el que está detrás de Chuck, entonces Alan no es el que está detrás de Chuck en la fila. Y, en este sentido, muchas relaciones son indiferentes porque no son ni transitivas ni intransitivas. Por ejemplo, si Ann está al lado de Beth, y Beth está al lado de Cath, no se sigue ni que Ann está al lado de Cath ni que no lo está. Podrían estar de pie en una fila o bien formando un pequeño corro.

Consideremos la relación espacial contenida en esta proposición:

Alan está detrás de Bill.

¿Qué se sigue? La respuesta depende de nuestro marco de referencia. Podemos utilizar un marco de referencia egocéntrico basado en nuestro punto de vista. Podemos decir: Tal como yo veo el mundo, Alan está detrás de Bill. Se sigue entonces que Bill está delante de Alan.

Por el contrario, podemos utilizar un marco de referencia objetivo independiente de nuestro punto de vista. Alan y Bill están de pie espalda con espalda, y por tanto Alan está detrás de Bill. Ya no se sigue que Bill está delante de Alan. Cuando decimos que estamos mirando hacia el oeste, estamos usando un marco de referencia objetivo que también es independiente del punto de vista. Los Guugu Yimithirr, el grupo australiano que describí en el Capítulo 2, tienen solamente un punto de vista objetivo respecto a las relaciones espaciales.

La ambigüedad respecto al marco de referencia recorre los términos ingleses para las relaciones espaciales. Pensamos en que muchos objetos tienen partes superiores e inferiores intrínsecas, partes delanteras y traseras intrínsecas, e incluso lados derechos e izquierdos intrínsecos. Una caja de madera cuadrada sin tapa tiene una parte de arriba y una parte de abajo intrínsecas, pero no una parte de delante y una parte de atrás intrínsecas, ni un lado derecho y un lado izquierdo intrínsecos. Ponga una tapa con bisagras en la parte de arriba, y mágicamente adquiere todas esas partes. Su parte delantera es el lado en el que se abre la tapa, y su lado derecho está a la derecha si nos colocamos frente a ella.

Las partes intrínsecas dependen de cómo nos orientamos respecto a los objetos –en este caso para abrir la tapa de la caja–. Los lados intrínsecos contrastan con los lados de los objetos que describimos desde nuestro punto de vista. El lado derecho intrínseco de la caja puede ser el lado izquierdo tal como la vemos. No asombra que los marineros utilicen una jerga especial para referirse a las partes intrínsecas de los barcos. “Atrás a toda máquina y girar a la izquierda”, sería confuso.

* * *

En el razonamiento relacional, nuestro propósito es inferir una nueva relación que se cumple entre los argumentos de las premisas. Podemos hacer estas inferencias sin basarnos en el recuerdo de una inferencia similar, y sin utilizar reglas de inferencia con un contenido particular. Por tanto, una vez más, la cuestión es si utilizamos la forma o el contenido. Consideremos esta inferencia, y en lo sucesivo supondré que el marco de referencia es egocéntrico:

El triángulo está a la derecha del círculo.

El círculo está a la derecha del rombo.

Por tanto, el triángulo está a la derecha del rombo.

¿Es válida? En el sentido de que su conclusión debe ser verdadera en el caso de que las premisas sean verdaderas, por supuesto que es válida. Desde un punto de vista formal, sin embargo, está haciendo desaparecer una premisa. Esta premisa establece que “a la derecha de” es una relación transitiva:

Si una primera entidad está a la derecha de una segunda entidad, y la segunda entidad está a la derecha de una tercera entidad, entonces la primera entidad está a la derecha de la tercera entidad.

Este principio es cierto para cualesquiera entidades en un esquema egocéntrico. Es axiomático. Los lógicos se refieren a estos axiomas como “postulados del significado” porque son postulados sobre los significados de los términos. Pueden parecer similares a reglas de inferencia con un contenido específico, pero las premisas –y los axiomas son premisas– no se deben confundir con reglas de inferencia. Utilizamos las reglas de inferencia para extraer conclusiones a partir de las premisas, no al revés. Las premisas son datos, a partir de los cuales las reglas de inferencia nos permiten sacar conclusiones.

Una vez que tenemos el postulado del significado para “a la derecha de”, podemos probar que la conclusión anterior se sigue de las premisas utilizando reglas formales de inferencia. La prueba exige la sustitución de las entidades concretas en el postulado del significado y luego una inferencia simple basada en una regla formal para las proposiciones condicionales. Los detalles de la prueba son contabilidad lógica y los he dejado para la Tabla 9.1. Las teorías del razonamiento que confían en reglas formales son claras y precisas. Pueden enfrentarse a cualquier aspecto del razonamiento relacional que se pueda capturar en postulados del significado. Y, a primera vista, sí parece que cualquier aspecto del razonamiento relacional puede explicarse de esa forma. Un importante lógico del siglo veinte, Yehoshua Bar-Hille, sugirió una forma útil de simplificar la contabilidad: tener un solo axioma multiuso para la transitividad, y poner una etiqueta en aquellas relaciones a las que es aplicable, como, por ejemplo, “a la derecha de”.

Tabla 9.1. Una prueba formal de una inferencia espacial transitiva

Las premisas son:

1. El triángulo está a la derecha del círculo.

2. El círculo está a la derecha del rombo.

3. Si una primera entidad está a la derecha de una segunda entidad, y la segunda entidad está a la derecha de una tercera entidad, entonces la primera entidad está a la derecha de la tercera entidad. (Un postulado del significado.)

La prueba se desarrolla como sigue:

4. Sustitúyanse las entidades concretas del postulado del significado para dar lugar a:

Si el triángulo está a la derecha del círculo y el círculo está a la derecha del rombo, entonces el triángulo está a la derecha del rombo.

5. Utilícese una regla formal para formar una conjunción de las premisas 1 y 2.

El triángulo está a la derecha del círculo y el círculo está a la derecha del rombo.

6. Utilícese una regla formal (si A entonces B; A; por tanto, B) con 4 y 5 para dar lugar a la conclusión:

Por tanto, el triángulo está a la derecha del rombo.

* * *

Un modelo mental de las relaciones espaciales es icónico e integra la información contenida en las premisas, y por tanto plasma relaciones nuevas que no corresponden a ninguna de las proposiciones utilizadas para su construcción. Cuando hacemos inferencias relacionales, podemos utilizar esas nuevas relaciones para formular conclusiones. A diferencia de un método basado en la forma, un sistema que utiliza modelos mentales no necesita axiomas para capturar las propiedades lógicas de las relaciones. La transitividad de una relación es una propiedad que surge de la construcción de modelos basados en los significados de las oraciones. Para ilustrar este punto, describiré un programa de ordenador que simula la teoría de modelos, pero, si no está usted interesado en el funcionamiento de los programas debería limitarse a echar solo un vistazo a esta explicación.

La entrada [input] del programa es una descripción espacial con o sin una conclusión dada, p. ej.:

El triángulo está a la derecha del círculo.

El círculo está a la derecha del rombo.

Por tanto, el triángulo está a la derecha del rombo.

El programa construye un modelo de la primera premisa. Utiliza el significado de el círculo para insertar un signo que representa al círculo en un modelo espacial tridimensional:

El significado de a la derecha de determina que el sistema busca en una dirección hacia la derecha del círculo, y por tanto el programa aumenta el eje izquierda-derecha a partir del círculo mientras mantiene constantes los valores en los otros dos ejes (arriba-abajo y delante- atrás). Luego utiliza el significado de el triángulo para insertar su representación en el primer espacio vacío del modelo:

El eje izquierda-derecha de este diagrama corresponde al eje espacial izquierda-derecha del modelo.

El programa puede buscar referentes en sus modelos espaciales. Por consiguiente, dada la segunda premisa:

El círculo está a la derecha del rombo.

descubre que el círculo ya está representado en su actual modelo de las premisas. Utiliza el significado de la oración para actualizar ese modelo. Por tanto, inserta en el modelo una representación del rombo en una posición adecuada:

Con la primera premisa, podemos partir del círculo para buscar, en la dirección que determina la relación, hasta encontrar un lugar para el triángulo. Pero, con la segunda premisa, este procedimiento natural no es factible. Como el programa, tenemos que buscar en la dirección opuesta a la que determina la relación, partiendo del círculo hasta encontrar un lugar para el rombo. Esta tarea debería ser un poco más difícil, y los datos muestran que lo es.

Si una premisa no se refiere a nada de lo que aparece en el modelo actual, el programa construye un nuevo modelo. Posteriormente, dada una premisa adecuada, puede integrar los dos modelos independientes en un único modelo. Esta situación también contribuye a añadir dificultad a nuestro razonamiento.

Dada la conclusión putativa en el ejemplo:

El triángulo está a la derecha del rombo.

el programa descubre que ambos referentes ya están presentes en su modelo actual. Comprueba si se cumple la relación apropiada entre ellos. Busca hacia la derecha del rombo hasta que encuentra el triángulo. La relación se cumple. A continuación comprueba si algún otro modelo de las premisas es un contraejemplo de la conclusión. No encuentra ninguno, y por tanto declara que la inferencia es válida. En caso de que una conclusión no se cumpla en el modelo actual, el programa comprueba si algún otro modelo de las premisas previas permite que la relación se cumpla. Si no es así, la proposición es inconsistente con lo anterior.

He descrito ya los principales procedimientos que utiliza el programa, y que están resumidos en la Tabla 9.2. Si nuestro sistema inferencial utiliza modelos, también los necesita. Estos procedimientos pueden producir inferencias válidas, funcionan sin un homúnculo y no utilizan postulados de significado. Ese tipo de axiomas no son necesarios, porque la validez de las inferencias surge de los significados de las relaciones, que determinan en qué dirección buscar en los modelos. La teoría de modelos postula que las inferencias relacionales las hacemos de forma similar.

Es fácil pasar por alto un punto. La búsqueda de contraejemplos por parte del programa funciona porque tiene acceso a las representaciones de los significados de las premisas. Sin estas representaciones, si el programa tuviera que cambiar un modelo, no tendría forma de comprobar si el resultado seguía siendo un modelo de las premisas. Cualquier sistema inferencial que construya modelos alternativos necesita por tanto un registro independiente de las premisas. Debe, o bien tener una memoria para sus significados, o bien ser capaz de volver a cada premisa y reinterpretarla.

La estrategia plasmada en el programa de razonamiento espacial es construir un único modelo cada vez. Cuando una descripción es consistente con más de una distribución, el programa construye el modelo que requiera el mínimo de trabajo. Una estrategia alternativa, que he implementado en un programa para el razonamiento sobre relaciones temporales, es intentar construir todos los diferentes modelos posibles. Aún existe otra estrategia que es representar las posibilidades alternativas dentro de un único modelo, utilizando un sistema para indicar qué posiciones de las entidades en el modelo no son seguras. Probablemente podemos desarrollar cualquiera de estas estrategias dependiendo de las particularidades de los problemas que abordemos. Esta noción de estrategia, como debería usted haber inferido, se refiere a la secuencia de pasos para la realización de inferencias que no se pueden llevar a cabo en una única operación mental, y en un capítulo posterior volveré a la cuestión de cómo desarrollamos estrategias inferenciales.

Tabla 9.2. Siete procedimientos para razonar utilizando modelos

1. Empezar un modelo nuevo. El procedimiento inserta un nuevo referente en el modelo de acuerdo con una premisa.

2. Actualizar un modelo con un nuevo referente en relación con un referente que ya existe.

3. Actualizar un modelo con una nueva propiedad o relación.

4. Reunir dos modelos independientes en uno de acuerdo con una relación entre referentes de ambos.

5. Verificar si una proposición es verdadera o falsa en los modelos.

6. Buscar un contraejemplo para refutar una proposición. Si la búsqueda falla, entonces la proposición se sigue de forma válida a partir de las proposiciones previas de la descripción.

7. Buscar un ejemplo para hacer verdadera una proposición. Si la búsqueda falla, entonces la proposición es inconsistente con las proposiciones previas.

* * *

Las identidades son un tipo especial de relación, pero también se pueden inferir a partir de modelos. Las inferencias se pueden dividir en dos clases principales, las que establecen una identidad y las que establecen una no identidad. Esta es una inferencia que establece una identidad:

Ernest es el señor Harrison.

El señor Harrison es el que bebe Bass.

Por tanto, Ernest es el que bebe Bass.

Y este es un ejemplo de una inferencia que establece una no identidad:

Ernest es el señor Harrison.

El señor Harrison no es el que bebe Guinness.

Por tanto, Ernest no es el que bebe Guinness.

Ambas inferencias son válidas, y muestran que podemos referirnos a sujetos tanto por un nombre como por una determinada descripción que distingue a un único sujeto, al menos en el contexto en que se utiliza.

Debido a que los modelos mentales son tan icónicos como es posible, cada entidad individual se representa en un modelo una sola vez. La proposición: Ernesto es el señor Harrison se representa como sigue:

Ernesto el señor Harrison

...

donde cada fila del diagrama representa un sujeto independiente. La elipsis permite la existencia de otros sujetos, y una nota mental a pie de página representa que ninguno de ellos es ni Ernest ni el señor Harrison. La proposición: el señor Harrison es el que bebe Bass actualiza el modelo:

Ernest el señor Harrison el que bebe Bass

...

Y este modelo da lugar a la conclusión:

Ernest es el que bebe Bass.

Por el contrario, la proposición negativa según la cual el señor Harrison no es el que bebe Guinness actualiza el modelo de esta forma:

Ernest el señor Harrison

el que bebe Guinness

...

y la nota a pie de página en el modelo implícito se actualiza para indicar que

ninguno de los sujetos que se representa en él bebe Guinness. El modelo representa dos sujetos: Ernest Harrison y el que bebe Guinness. Da lugar a la conclusión:

Ernest no bebe Guinness.

Tabla 9.3. La solución para un problema de identidad

Ernest, el señor Littlewood y otros dos hombres prefieren cada uno una determinada cerveza entre cuatro tipos diferentes, incluida la Guinness. Los cuatro hombres se identifican únicamente como el que bebe Bass, el señor Hardy, y dos que no beben Adnams: John y el señor Russell.

Ni Bertie ni el señor Harrison que es el amigo de Godfrey beben Harp.

Bertie que es el vecino del señor Hardy no bebe Bass.

Para resolver el problema situemos a los cuatro sujetos únicos en una tabla:

	1	2	3	4
Nombre			John	

Apellido		Hardy	Russell
Cerveza preferida	Bass		

Dado que ni John ni Russell beben Adnams, y la persona 1 bebe Bass, se sigue que Hardy bebe Adnams. Bertie es el vecino de Hardy y por tanto no puede ser Hardy. Bertie tampoco es John ni el que bebe Bass, y por tanto debe ser Russell. Como Bertie no bebe Harp, debe beberla John. Por consiguiente, Bertie debe beber Guinness. Harrison no bebe Harp, y por tanto debe ser el que bebe Bass. En consecuencia, Littlewood es John. Harrison es el amigo de Godfrey, y por tanto Godfrey es Hardy, y Harrison es Ernest. La solución según eso es:

Nombre	Ernest	Godfrey	John	Bertie
Apellido	Harrison	Hardy	Littlewood	Russell
Cerveza preferida	Bass	Adnams	Harp	Guinness

Los juegos de lógica de las revistas y de los test psicológicos están basados a menudo en descripciones de sujetos y de sus propiedades exclusivas (ver la Tabla 9.3). Las dos inferencias centrales en estos juegos se refieren una vez más a la identidad y a la no identidad. Si un sujeto es no idéntico a todos los miembros de un conjunto excepto uno, entonces el sujeto es idéntico a esa excepción. De la misma forma, si en un conjunto de relaciones exclusivas una le corresponde a un sujeto, p. ej., él es el que bebe Guiness, se sigue que Guinnes no es la cerveza de ningún otro sujeto. Una forma de establecer que dos expresiones se refieren a diferentes sujetos es partir de una relación que no

puede establecerse entre una entidad y ella misma, p. ej., dado que Bertie es el vecino de Hardy, se deduce que Bertie no es Hardy. Uno no puede ser su propio vecino. La Tabla 9.3 presenta un juego de identidad, y da la solución basándose en la implementación de estos principios que hace mi programa de ordenador.

* * *

Los límites de la iconicidad pueden ser mayores de lo que he descrito hasta ahora en este libro. Para entender por qué, considere este modelo y pregúntese a sí mismo qué representa:

Anteriormente lo utilicé para representar las relaciones espaciales entre las tres formas. Pero, solo con el modelo, usted no tiene medios para saberlo. Podría representar que están presentes las tres formas, o que aparecen una detrás de otra en ese orden: rombo, círculo, triángulo. En un modelo espacial los ejes representan dimensiones espaciales; en el modelo de la mera presencia de las formas los ejes no representan nada; en un modelo temporal, el eje principal representa el tiempo. El cerebro podría mantener sistemas independientes dedicados a estos diversos tipos de modelos. Pero una secuencia de sucesos espaciales puede desarrollarse en el tiempo, y por tanto el sistema espacial tendría que incrustarse en el sistema temporal. A veces una secuencia temporal puede afectar a la presencia de objetos, de forma que estos tipos de modelos tendrían también que incrustarse en el sistema temporal. Para empeorar las cosas, algunas veces podemos representar una secuencia temporal en un modelo espacial estático, como por ejemplo un gráfico.

Quizá una posibilidad más plausible es similar a lo que sucede cuando un coronel retirado nos explica la batalla de El Alamein sobre la mesa de la cena. Utiliza los botes de la sal y la pimienta y distribuye piezas de la cubertería para representar las disposiciones de los ejércitos de Rommel y Montgomery. El

problema es: ¿cómo sabemos nosotros que el bote de la sal representa a la 1ª División Acorazada de Montgomery? La respuesta es que nos lo dice el coronel, y nosotros y él podemos mantenerlo presente. De la misma manera, cuando el sistema de razonamiento construye un modelo de una disposición espacial, puede representar que los ejes tienen una interpretación espacial. Como usted ya sabe, el sistema de razonamiento necesita una representación independiente de los significados de las oraciones, de manera que cuando modifica un modelo para buscar un contraejemplo, puede comprobar que el modelo resultante sigue representando las premisas. Por tanto, los significados de las oraciones podrían establecer cómo hay que interpretar el modelo. En este caso, lo que un modelo representa depende tanto del modelo como del significado de las oraciones utilizadas para construirlo.

* * *

Las pruebas corroboran la utilización de modelos en el razonamiento. Los participantes en experimentos informan de que imaginan distribuciones espaciales. A menudo hacen gestos con las manos que sugieren que tienen en mente un modelo espacial. De la misma forma, si tienen papel y lápiz, dibujan diagramas. Sin embargo, las pruebas no descartan la posibilidad de que en el fondo los procesos inferenciales inconscientes estén guiados por la forma más que por el contenido. ¿Existe, quizá, un experimento crucial que podría mostrarnos que la gente se basa en un método en lugar de en el otro? Describiré tres.

El primer experimento utilizaba descripciones de las distribuciones espaciales bidimensionales de objetos domésticos que podríamos encontrar sobre una mesa. Un tipo de problemas era, por ejemplo:

La taza está a la derecha del plato.

La cuchara está a la izquierda del plato.

El cuchillo está delante de la cuchara.

El platillo está delante de la taza.

¿Cuál es la relación entre el cuchillo y el platillo?

Si utilizamos reglas formales para obtener la respuesta, tenemos que proceder mediante dos pasos principales. Tenemos que utilizar postulados del significado y las dos primeras premisas para inferir la relación entre la cuchara y la taza:

La cuchara está a la izquierda de la taza.

Y luego tenemos que utilizar un postulado del significado “bidimensional” para obtener con esta conclusión y las dos premisas restantes la respuesta a la pregunta:

El cuchillo está a la izquierda del platillo.

Por consiguiente, primero extraemos una conclusión transitiva y luego la utilizamos para extraer una conclusión sobre las entidades de la pregunta.

Un segundo tipo de problema era, por ejemplo:

El plato está a la derecha de la taza.

La cuchara está a la izquierda del plato.

El cuchillo está delante de la cuchara.

El platillo está delante del plato.

¿Cuál es la relación entre el cuchillo y el platillo?

En este caso no se necesita la inferencia transitiva inicial. La segunda premisa afirma que la cuchara está a la izquierda del plato, y el postulado "bidimensional" y esta conclusión dan lugar a la respuesta de que el cuchillo está a la izquierda del platillo. Este problema debería ser por tanto más fácil que el primero.

La teoría de modelos hace la predicción opuesta. Las premisas del primer problema dan lugar a un modelo bidimensional de la distribución espacial de los objetos mostrado en este diagrama como si lo viéramos desde arriba:

Cuchara plato taza

cuchillo platillo

Este modelo da lugar a la conclusión de que el cuchillo está a la izquierda del

platillo. Las premisas del segundo problema, sin embargo, son consistentes con dos disposiciones distintas, porque no fijan la relación entre la cuchara y la taza:

cuchara taza plato taza
cuchara plato

cuchillo platillo
cuchillo platillo

Ambas disposiciones, sin embargo, apoyan la conclusión de que el cuchillo está a la izquierda del platillo, pero los dos modelos –o un único modelo representando las posiciones alternativas de la taza– deberían ser más difíciles de mantener presentes, y por tanto este problema debería ser más difícil que el primero.

Puede usted haberse dado cuenta de que la primera premisa del segundo problema es irrelevante para la inferencia. La premisa puede provocar una disrupción en el razonamiento, y por ello conducir a que este problema tenga una mayor dificultad. Para comprobar esta posibilidad, el experimento utilizaba también problemas que dan lugar a un único modelo, pero que tienen una primera premisa irrelevante. Si la aparición de una premisa irrelevante es crítica, estos problemas deberían ser más difíciles que los del primer tipo, aunque ambos requieran un único modelo. Pero, si la teoría de modelos es correcta, la premisa irrelevante no debería provocar diferencias. Los participantes no tuvieron dificultades con las premisas irrelevantes, y tuvieron un setenta por ciento de respuestas correctas en los problemas con una única disposición pero solo un cuarenta y seis por ciento de respuestas correctas en los problemas con más de una disposición. Las conclusiones correctas fueron también más rápidas para las disposiciones únicas (una media de 3,1 segundos) que para las disposiciones múltiples (3,6 segundos). Otros estudios han replicado estos resultados. Han mostrado también el mismo efecto del número de posibilidades respecto a relaciones temporales como antes y después, y respecto a relaciones más abstractas como estudiar más que y copiar de, que deberían tender a evocar modelos abstractos. Incluso cuando todos los problemas tienen una premisa

irrelevante, los problemas que dan lugar a un modelo único son más fáciles que los que dan lugar a modelos múltiples.

Un segundo experimento examinaba el razonamiento sobre relaciones temporales, pero dependía de los verbos y los tiempos verbales para establecer el orden de los sucesos. Este es un problema típico del experimento:

John ha limpiado la casa.

John se está duchando.

John va a leer el periódico.

Mary siempre lava los platos cuando John limpia la casa.

Mary siempre se toma el café cuando John lee el periódico.

¿Cuál es la relación entre que Mary lave los platos y se tome el café?

El tiempo de los verbos y los diversos auxiliares, como “ha”, establece el orden de las tres acciones de John: limpia la casa, se ducha y luego lee el periódico. Las premisas cuarta y quinta enlazan las dos acciones de Mary con las de John, y por tanto podemos construir un modelo temporal de toda la descripción. El siguiente diagrama dibuja el modelo, en el que el tiempo es el eje de izquierda a derecha, y cada línea muestra la secuencia de acciones de uno de los protagonistas:

John: limpia la casa se ducha lee el periódico

Mary: lava los platos se toma el café

Se sigue que Mary lava los platos antes de tomarse el café. Un problema consistente con dos órdenes temporales diferentes tiene una segunda premisa diferente:

John se ha duchado.

En este caso, hay dos órdenes posibles para las dos primeras acciones de John, porque no puede hacer las dos cosas al mismo tiempo. El experimento mostraba que se producían respuestas correctas con mayor frecuencia para los problemas que dan lugar a un solo modelo mental que para los problemas que dan lugar a modelos múltiples. La dificultad de una descripción formal reside en la elaboración de postulados del significado que hagan posibles estas inferencias. Es difícil, porque las inferencias dependen de verbos auxiliares, tiempo verbal y aspecto verbal. Los postulados tendrían que mezclar todos estos aspectos, p. ej.:

Si una oración contiene “ha” seguido de un verbo en pasado y describe un suceso, y si una segunda oración contiene “va a” antes de un infinitivo y describe un suceso, entonces el suceso descrito en la primera oración se produce antes que el suceso descrito en la segunda oración.

Este postulado viola el principio de que los postulados del significado se refieren a las propiedades lógicas de las palabras.

Un tercer experimento demostró una dificultad aún mayor para los postulados del significado. Examinaba inferencias del tipo de:

Alice es pariente consanguíneo de Brian.

Brian es pariente consanguíneo de Charlie.

¿Qué se sigue?

Puede estar usted tentado de inferir que Alice es pariente consanguíneo de Charlie. Quizá piense usted en un conjunto de hermanos o una línea de descendientes. Sin embargo, hay contraejemplos para la conclusión. Supongamos, por ejemplo, que Alice es la madre de Brian y que Charlie es su padre. Alice está relacionada con Brian, y él está relacionado con Charlie, pero su madre y su padre probablemente no son parientes consanguíneos. Estas inferencias “pseudotransitivas” se basan en relaciones que no son ni transitivas ni intransitivas, pero que dan lugar a modelos de típicas situaciones en las que se cumple una conclusión transitiva.

Cuando razonamos utilizamos nuestro conocimiento. Dadas estas premisas:

Armstrong adelantó a Hamilton.

Ullrich adelantó a Armstrong.

es probable que infiramos que Ullrich adelantó a Hamilton. Pero, hay un contraejemplo. Supongamos, por ejemplo, que las premisas se refieren a una

contrarreloj en el Tour de Francia. Los corredores salen de uno en uno en intervalos de dos minutos. Armstrong alcanza a Hamilton, y le adelanta. Hamilton abandona la carrera en ese punto. Y luego Ullrich, que salió después que Armstrong le alcanza y le adelanta. Ambas premisas son verdaderas, pero debido a esos sucesos extraordinarios, la conclusión transitiva es falsa. Ullrich nunca adelantó a Hamilton, porque Hamilton no siguió en la carrera. Por consiguiente, podría haber estado esperando en la meta la llegada de Ullrich.

Hacemos inferencias pseudotransitivas porque las premisas suscitan un modelo de una situación típica en la que la conclusión transitiva funciona. La teoría de modelos, sin embargo, predice que la forma de bloquear estas inferencias es hacer que los participantes busquen contraejemplos con más ahínco. Otro experimento valoraba a los participantes en dos situaciones diferentes. En una situación, llevaban a cabo los problemas pseudotransitivos originales. En la otra situación, cada problema pseudotransitivo se presentaba con una pista verbal diseñada para ayudar a los participantes a pensar en un caso atípico. Por ejemplo, el problema sobre los “parientes consanguíneos” venía precedido por la pista: las personas pueden estar emparentadas por consanguinidad o por matrimonio. Cada participante encontraba un determinado problema una sola vez, en una de las dos situaciones. Las pistas tenían un fuerte efecto sobre el razonamiento. Sin ellas, los participantes extrajeron conclusiones pseudotransitivas en el sesenta y dos por ciento de las ocasiones; pero con ellas, este porcentaje cayó hasta el cuarenta y uno por ciento.

Si utilizamos reglas formales para razonar, dependemos de un postulado del significado que captura la transitividad de la relación. En el esquema de Bar-Hillel, cada relación que considera dos argumentos se etiqueta como transitiva, intransitiva o ninguna de las dos cosas. Por tanto, ¿qué tipo de relaciones deberían etiquetarse como transitivas? Podría suponerse que entre las candidatas adecuadas estarían relaciones como por ejemplo: “más alto que”. Pero considere este problema:

Cate es más alta que Belle.

Belle era más alta que Alice (hasta que Alice bebió la poción mágica).

¿Quién es más alta?

El cambio en el tiempo verbal y la poción mágica ya no garantizan la transitividad. La esencia de este ejemplo es que ningún término relacional, ni siquiera “más alto que”, puede clasificarse como transitivo en todos los casos. La transitividad depende de la significación de la proposición como un todo, lo cual a su vez depende del contexto. Por consiguiente, las propiedades lógicas de los términos relacionales en el lenguaje natural no se pueden capturar en postulados del significado. Afortunadamente, no necesitamos postulados del significado si razonamos utilizando modelos. Consideramos el significado de las premisas, tenemos en cuenta el contexto y nuestro conocimiento, y luego imaginamos todas las posibilidades compatibles con esa información. Si una conclusión transitiva se cumple en todas esas posibilidades, entonces la inferencia es válida.

El poeta francés Paul Valéry sugirió una vez que la cuestión siguiente podría ser una buena prueba de talento matemático:

Si Peter se parece a Paul, y Paul se parece a James, ¿se parece James a Peter?

Aquellos que dicen “sí” a la primera son los dotados. Sus esperanzas en la prueba se hicieron añicos cuando uno de los mejores matemáticos que conocía empezó a meditar largamente. No hay que preocuparse, porque es discutible si la relación es transitiva. Podría decir que Marx se parece a Lenin, y que Lenin se parece a Stalin, pero no que Marx se parece a Stalin. Quizá una prueba mejor sería pedir a los futuros matemáticos que resolvieran este problema:

Yo estaba casado con una viuda que tenía una hija ya mayor. Mi padre se enamoró de ella, y pronto los dos se casaron. Esto convirtió a mi padre en mi yerno y ahora mi hija era mi madre, porque era la mujer de mi padre. Para complicar el asunto, pronto me convertí en padre de un saludable bebé. Mi bebé

se convirtió así en un cuñado de mi padre, y por tanto se convirtió en mi tío. Si era mi tío, entonces eso le hacía también hermano de la hija mayor de la viuda, quien, evidentemente, era mi madrastra. La mujer de mi padre tuvo entonces un hijo que se convirtió en mi nieto, porque era el hijo de mi hija. Mi mujer es ahora la madre de mi madre, y aunque es mi mujer, es también mi abuela. Ahora, si mi mujer es mi abuela, yo soy su nieto, y me he convertido en el caso más extraño que usted haya visto nunca. ¿Qué soy?

* * *

La teoría de modelos postula que nuestro razonamiento depende de la significación de las proposiciones. La transitividad de las relaciones es una propiedad que surge de la construcción de modelos icónicos. Estos modelos representan muchas relaciones que no se establecen en el discurso, y por tanto nuestra tarea es encontrar una relación útil y formular una conclusión que la incluya. Los datos experimentales corroboran esta descripción. Los programas de ordenador están produciendo modelos que muestran que se puede hacer todo mediante procesos simples. No se necesita magia; no se necesitan homúnculos, no se necesitan postulados del significado.

Notas al pie

[1] Nota de la trad.: En inglés este título incluye un juego de palabras, ya que relation significa no solo relación sino también parentesco/parientes, a lo largo del capítulo utilizaré en cada caso el término que mejor se adapte al contexto.

CAPÍTULO 10

Los silogismos y el razonamiento sobre propiedades

Supongamos que acabo de llegar de una exposición de arte, y usted me pregunta: "¿Qué tipo de artistas son?" Respondo:

La mayoría son pintores y escultores.

Y usted dice:

Así que la mayoría son escultores y pintores.

Usted ha hecho una inferencia sobre dos propiedades –la propiedad de ser un pintor y la propiedad de ser un escultor–. Y la inferencia es trivial porque no hace más que reordenar esas dos propiedades. Pero también es sorprendente, porque ninguna de las actuales teorías psicológicas sobre el razonamiento basadas en reglas formales puede explicar cómo la ha hecho usted. El problema que se les plantea es el cuantificador, "la mayoría". Queda fuera del campo de sus reglas formales, que se refieren solo a cuantificadores estándar como "cada uno" y "algunos". Los lógicos se dieron cuenta hace mucho tiempo de que "la mayoría" y otros cuantificadores, como "más de la mitad", son diferentes de los cuantificadores estándar, y exigen una lógica más potente. Esta lógica de alta potencia proporciona una manera uniforme de manejar todos los cuantificadores. En esencia, considera que una proposición como por ejemplo:

La mayoría de los artistas son pintores y escultores.

tiene el significado de que el conjunto de artistas que son pintores y escultores es mayor que el conjunto de artistas que no son ambas cosas: pintores y escultores. Las teorías psicológicas basadas en reglas formales todavía no se han formulado de ese modo.

Voy a intentar llegar a un acuerdo con las teorías formales. Supongamos que yo hubiera contestado a su pregunta sobre los artistas:

Cada uno es pintor y escultor.

Y usted infiriera:

Así que cada uno es escultor y pintor.

Las teorías formales pueden explicar cómo hace esta inferencia. Utiliza una regla para extraer el cuantificador y reemplazarlo por un sujeto arbitrario del conjunto considerado, p. ej., Fred es pintor y Fred es escultor. Utiliza otra regla para separar la primera cláusula de esta conjunción, y luego la segunda, y otra regla para unirlas en el orden opuesto. Finalmente, utiliza otra regla más para reemplazar al sujeto arbitrario por el cuantificador:

Cada uno es escultor y pintor.

Por consiguiente, tiene usted que hacer una secuencia de cinco inferencias independientes para extraer su conclusión. Incluso los incondicionales de las

reglas formales serían escépticos ante el hecho de que hacer una inferencia tan trivial suponga tantos pasos.

Un remedio obvio es postular una regla formal que lleve a cabo la inferencia en un único paso. Este remedio exige un coste: es ad hoc, y aún peor, aumenta el número de reglas de la teoría, y por tanto, al buscar una prueba, en cualquier nivel de la prueba hay más reglas que comprobar. Pero hay otro problema más general. El primer paso al usar las reglas formales es extraer las formas lógicas de las premisas para emparejarlas con las reglas apropiadas. La recuperación de esta representación es similar a la traducción de una oración a una fórmula lógica. Pero este problema es tan difícil que no existe ningún programa de ordenador que traduzca oraciones inglesas a sus formas lógicas. Y por tanto no existe ningún método automático para utilizar la lógica para evaluar argumentos expresados en lengua vernácula. Leibniz soñó con un método así hace más de trescientos años. Su sueño todavía no se ha cumplido y eso, como dijo otro lógico, Bar-Hillel, es un escándalo de la existencia humana.

* * *

Una forma de hacer las dos inferencias anteriores es contratar a un grupo de actores para que desempeñen los distintos papeles. Podrían formar un cuadro para representar la proposición, la mayoría de los artistas son pintores y escultores. Podrían sostener pinceles y paletas para representar que son pintores y mazos y cinceles para representar que son escultores. Esta analogía revela una asunción central en la teoría de modelos. Un modelo mental representa un conjunto de entidades como un conjunto icónico de referentes. Es la representación abstracta que está detrás de cualquier imagen de los actores que pueda usted formar en su imaginación. Por consiguiente, el modelo mental de la proposición es como el cuadro de los actores, pero no representa los detalles de su apariencia o cómo se ven desde un determinado punto de vista. Utiliza un número arbitrario de referentes para representar al conjunto de sujetos:

artista pintor escultor

artista pintor escultor

artista pintor escultor

artista

Cada fila del diagrama corresponde a un único sujeto, y por tanto la primera fila representa a un sujeto que es artista, pintor y escultor. Como puede usted ver, el modelo representa la proposición de que la mayoría de los artistas son pintores y escultores, porque tres de cuatro lo son. Este modelo también se puede describir como:

La mayoría de los artistas son escultores y pintores.

Por consiguiente, la conclusión que usted extrajo es un resultado inmediato del modelo y no se necesita la complicada secuencia de inferencias formales que describí anteriormente.

Un modelo similar representa la proposición de que cada uno es pintor y escultor:

artista pintor escultor

artista pintor escultor

artista pintor escultor

artista pintor escultor

También da lugar a la conclusión inmediata con las dos propiedades en orden opuesto:

Cada uno es escultor y pintor.

Sin embargo, me estoy adelantando a mí mismo ligeramente, porque, ¿cómo sabemos que el modelo anterior, o incluso el grupo de actores, representa a todo el conjunto relevante, en lugar de representar solamente, pongamos por caso, a cuatro de los sujetos del conjunto? En otras palabras, ¿cómo sabemos que el modelo representa un conjunto completo?

Ningún modelo icónico puede mostrar que representa un conjunto completo. Esta información está en el significado de la oración representada por el modelo, y anteriormente sostuve que el razonamiento basado en modelos depende también de una representación de ese significado. Esta representación es necesaria con el fin de que cualquier modificación de un modelo se pueda revisar para asegurarse de que el resultado sigue siendo un modelo de las premisas. Ya hemos encontrado otra razón para tener acceso al significado de las premisas.

En los programas de ordenador que simulan la teoría, a veces he utilizado una nota a pie de página especial para señalar la representación de un conjunto completo. Por ejemplo, una proposición como:

Cada pintor es escultor

tiene el modelo

pintor escultor

pintor escultor

pintor escultor

...

La elipsis representa a otros sujetos en la situación descrita, pero tiene una nota a pie de página según la cual entre ellos no aparecen pintores. Por tanto, el modelo representa el conjunto de pintores completo.

¿Cómo sabemos que el modelo no significa que hay solo tres pintores en el conjunto? El problema es distinguir entre propiedades accidentales de una representación, tales como el número de pintores del modelo anterior, y propiedades esenciales, tales como el hecho de que representa un conjunto completo. Los filósofos lucharon durante mucho tiempo con este tipo de problema. ¿Cómo puede, por ejemplo, un único diagrama de un triángulo representar en una demostración geométrica a cualquier triángulo? Sus ángulos concretos son accidentales, pero el hecho de que es un triángulo es esencial. Kant consideró que la tarea era imposible, y por tanto los conceptos se tenían que representar como esquemas abstractos. Wittgenstein sostuvo por el contrario que un único objeto, como por ejemplo una hoja, podría representar a cualquier hoja, con tal de que se utilizara de la forma adecuada. No se paró a explicar cuál

era la forma adecuada. Sin embargo, David Hume, el gran filósofo de la Ilustración escocesa, había anticipado la solución:

... despúes que el espíritu ha producido una idea individual, sobre la que razonamos, el hábito que la acompaña y es despertado por el término general o abstracto sugiere rápidamente otra idea individual si por casualidad hacemos un razonamiento que no concuerda con aquella. Así, si mencionáramos la palabra triángulo, y formásemos la idea de un equilátero determinado para corresponder a aquella y afirmásemos después que los tres ángulos de un triángulo son iguales entre sí, las otras ideas individuales de un escaleno y un isósceles, que hemos omitido al principio, se nos presentan inmediatamente y nos hacen percibir la falsedad de esta proposición...

En resumen, los especialistas en geometría saben que se trata de que el diagrama represente un triángulo arbitrario, y por tanto pueden pensar en otro tipo de triángulo para refutar una conclusión errónea que se cumple solo para su diagrama inicial. De la misma forma, sabemos que la elección de tres pintores en el modelo anterior era arbitraria y por eso podemos cambiar el número.

* * *

Los procesos subyacentes en el razonamiento sobre conjuntos son sencillos y son un reflejo de los utilizados en el razonamiento sobre identidades y no identidades. Son fundamentales dos tipos de inferencias. Inferimos que las entidades pertenecen a un conjunto o inferimos que no pertenecen a un conjunto. La Tabla 10.1 ilustra ambos tipos de inferencia. Ambas giran sobre una propiedad, o la falta de esta, que se cumple para un conjunto completo. Dada una premisa que afirma esa propiedad, otra premisa se puede basar en un cuantificador estándar, como por ejemplo “algunos” o “todos”, o en cuantificadores que no son estándar, como “la mayoría” o “más de la mitad”. Por ejemplo, esta inferencia es válida:

La mayoría de los artistas son escultores.

Cada escultor es pintor.

Por tanto, la mayoría de los artistas son pintores.

Tabla 10.1. Las dos inferencias fundamentales sobre conjuntos

1. El primer tipo de inferencia establece que un sujeto pertenece a un conjunto, p. ej.:

Stella es escultora.

Todo escultor es pintor.

Por tanto, Stella es pintora.

Las premisas dan lugar al modelo mental:

Stella escultor pintor

escultor pintor

escultor pintor

...

No importa con qué escultor se identifique a Stella, se sigue que Stella es pintora, porque el modelo representa el conjunto de escultores completo. Por consiguiente, cuando las entidades están en un conjunto, heredan todas las propiedades que se cumplen para el conjunto completo.

2. El segundo tipo de inferencia establece que un sujeto no pertenece a un conjunto, p. ej.:

Cristo no es pintor.

Todo escultor es pintor.

Por tanto, Cristo no es escultor.

Las premisas dan lugar al modelo:

Cristo

escultor pintor

escultor pintor

escultor pintor

...

Se sigue que Cristo no es escultor. Por consiguiente, cuando las entidades no tienen una propiedad que se cumple para un conjunto completo, no pertenecen al conjunto.

Imagino que Aristóteles era consciente del principio. Ciertamente, sus seguidores escolásticos de la Edad Media lo eran –se referían a los términos que implicaban un conjunto completo como “distribuidos”.

Una vez que podemos inferir si los sujetos pertenecen o no a un conjunto, podemos hacer frente a todas las inferencias sobre propiedades, porque las propiedades definen los conjuntos. Los procesos que nos capacitan para combinar la información de las premisas independientes en modelos únicos son los que ya he descrito para el razonamiento relacional y proposicional, pero preparados para ocuparse de los sujetos que pertenecen a conjuntos. Forman modelos, los actualizan, formulan conclusiones o las verifican, y buscan modelos alternativos.

* * *

Si algunos de los pintores son escultores, entonces se sigue que algunos de los escultores son pintores. Pero, si más de la mitad de los pintores son escultores, ¿se sigue que más de la mitad de los escultores son pintores? Un principio general en la construcción de modelos es la parsimonia. Tratamos de representar tan pocas clases de sujetos distintos como sea posible. De ahí que es probable que representemos la última premisa como sigue:

pintor escultor

pintor escultor

pintor

...

La nota a pie de página en el modelo implícito establece que los sujetos que representa no son pintores. Si suponemos erróneamente que el modelo representa también el conjunto de escultores completo, inferiremos que se sigue la conclusión. Pero, por supuesto, podría haber otros escultores, p. ej.:

pintor escultor

pintor escultor

pintor

escultor

escultor

...

Este modelo es un contraejemplo para la conclusión. Los experimentos sugieren que los sujetos sí extraen ese tipo de conclusiones no válidas.

* * *

Supongamos que le cuento a usted algo más sobre la exposición de arte que he visitado:

Algunos de los artistas son diseñadores, y todos los diseñadores son minimalistas.

Usted puede inferir que algunos de los artistas son minimalistas. La inferencia es un ejemplo de silogismo. Aristóteles diseñó una lógica para los silogismos, que

son inferencias sobre propiedades basadas en dos premisas cuantificadas. Restringió los cuantificadores a “todos”, “algunos” y sus negaciones. La Tabla 10.2 describe los silogismos. Son parte de la lógica de propiedades, que a su vez es parte de la lógica moderna. Pero gran parte de la lógica tradicional se basaba en los silogismos de Aristóteles. Cada teólogo medieval que se preciara era experto en su utilización. La BBC emitió una vez un debate monástico expresado solo en silogismos, y los debatientes mostraron una gran habilidad para presentar los argumentos ajustándose a su rígida forma.

Tabla 10.2. Silogismos

Un silogismo categórico, o silogismo para abreviar, es un argumento formado por dos premisas y una conclusión. Cada una de esas tres oraciones tiene una de las siguientes “formas”:

Todos los ____ son ____

Algunos ____ son ____

Ningún ____ es un ____

Algunos ____ no son ____

Por ejemplo, las siguientes premisas dan lugar a una conclusión silogística: Algunos artistas son diseñadores; todos los diseñadores son minimalistas. Los

términos de las premisas, p. ej., artistas, diseñadores y minimalistas, se pueden distribuir en cuatro “figuras” diferentes:

A-B	B-A	A-B	B-A
B-C	C-B	C-B	B-C

Hay por tanto sesenta y cuatro posibles tipos de premisas silogísticas (4 formas x 4 formas x 4 figuras). Los lógicos medievales solo permitían conclusiones con la forma C-A. Los psicólogos se dieron cuenta de que también podrían tener la forma: A-C.

De ahí que haya 512 formas posibles de silogismos completos (sesenta y cuatro tipos de premisas x 4 formas de la conclusión x 2 figuras de la conclusión).

Los lógicos medievales identificaban los silogismos válidos utilizando reglas como por ejemplo:

1. Para que las premisas den lugar a una conclusión válida, el término que aparece en las dos premisas debe estar “distribuido” al menos en una de ellas (es decir, referirse a todo un conjunto).

2. Ningún término puede estar distribuido en la conclusión a menos que esté distribuido en las premisas.

3. Si ambas premisas contienen “algunos”, entonces no se sigue ninguna conclusión categórica.

4. Si ambas premisas son afirmativas, la conclusión es afirmativa; en caso contrario es negativa.

En la lógica moderna, una oración como, “Se disparará a cualquier desertor”, no implica que haya desertores. En cambio, Aristóteles y sus seguidores de la Edad Media asumían que todos los términos de un silogismo se referían a entidades existentes. Los psicólogos dicen a los participantes en sus experimentos que existen sujetos de ambos tipos. En este caso, de las sesenta y cuatro posibles formas de las premisas veintisiete dan lugar a conclusiones categóricas válidas.

▪

Bertrand Russell, uno de los fundadores de la lógica moderna, no tenía tiempo para silogismos, y una vez sostuvo que el único uso que alguna vez había oído que tuvieran era el que les dio un filósofo alemán que se asombró con los anuncios de un número humorístico del periódico, Mind. Al final el filósofo se argumentó a sí mismo: cada cosa de este libro es una broma; los anuncios están en este libro; por tanto, los anuncios son una broma. Los silogismos son objeto de muchas otras bromas. Woody Allen en su película Love and Death argumentaba: “Todos los hombres son mortales. Sócrates era mortal. Por tanto, todos los hombres son Sócrates. Lo que significa que todos los hombres son homosexuales.” No hay broma como una broma antigua. Platón tenía un argumento sofista: Ese perro es padre; ese perro es tuyo; por tanto, ese perro es tu padre.

Sí que hacemos inferencias sobre propiedades, pero normalmente sobre propiedades de los sujetos, como en las inferencias sobre Stella y Cristo. Y los silogismos sí se producen en la vida fuera de los monasterios, pero tienden a estar escondidos en el lenguaje informal, y a menudo dan por hecha una premisa no enunciada, p. ej., todos los aparatos eléctricos son británicos, y por eso no funcionarán en Estados Unidos. Una vez, después de que yo diera una charla sobre silogismos, un distinguido psicólogo animal me dijo: el problema con los silogismos es que son artificiales. Le dije: supongo que usted piensa que los psicólogos no deberían estudiar problemas artificiales. Él contestó: sí. Le dije: por tanto, usted infiere que los psicólogos no deberían estudiar los silogismos. Sí, contestó él. Y luego se rio. No necesité explicarle que entre los dos acabábamos de construir un silogismo, aunque no tuviera completo su traje formal.

* * *

La mayoría de los estudios psicológicos sobre el razonamiento respecto a las propiedades han examinado los silogismos. La razón, sospecho, es la funesta influencia de la lógica tradicional. Los estudios iniciales presentaban casi siempre conclusiones para evaluar, o de forma individual o en un test de opción múltiple. Los resultados llevaron a algunos investigadores a sostener que a menudo no razonamos, sino que elegimos una respuesta que encaje con la "atmósfera" de las premisas. Por ejemplo, si ambas premisas contienen el cuantificador "todos", se supone que tenemos más probabilidad de aceptar, solo en base a eso, una conclusión que contenga "todos". Las diversas versiones de esta opinión tienen muchos partidarios aún hoy. Por supuesto, es imposible mostrar que ninguno de los participantes en un experimento hayan sido nunca afectados por la atmósfera. El mismo tipo de efecto lleva a errores gramaticales por descuido, p. ej., en la oración anterior debería haber aparecido "haya" en lugar de "hayan". Pero sospecho que el efecto atmósfera se ha exagerado debido a una desafortunada coincidencia: la mayoría de las conclusiones válidas sí encajan con la atmósfera de las premisas. Cuando extraemos conclusiones con nuestras propias palabras, a menudo respondemos, "No se sigue nada". La hipótesis de la atmósfera no puede explicar esta respuesta, porque siempre hay una conclusión putativa que se ajusta a la atmósfera de las premisas. Además, damos esta respuesta incluso cuando existe una conclusión válida e incluso una que encaja con la atmósfera de las premisas.

Un experimento sobre silogismos utilizó premisas como, por ejemplo, "Solo los solitarios son valientes". Incluso cuando ambas premisas contenían el cuantificador "solo", los participantes tenían una probabilidad tres veces mayor de extraer una conclusión que contuviera "todos" que una que contuviera "solo". Lo que me convence de que quienes no han recibido enseñanzas de lógica hacen deducciones mediante silogismos es la popularidad universal de los Sudokus, que giran en torno a relaciones cuantificadas más complejas y que no se pueden resolver por ninguna vía que no sea la deducción. Volveré a ellos posteriormente en otro capítulo del libro. La creencia en la hipótesis de la atmósfera podría ser el mejor ejemplo del fenómeno que pretende explicar –un error al razonar.

Los estudios sobre silogismos dan dos resultados sólidos. El primero es que los sujetos se diferencian en habilidad. En una universidad con unos estrictos

criterios de selección para la admisión de estudiantes, un estudio mostró que la exactitud de los participantes variaba entre un cincuenta y cinco por ciento y un ochenta y cinco por ciento de respuestas correctas. Un estudio similar en una universidad con admisión libre mostró que la exactitud de los participantes variaba entre un treinta y siete por ciento y un cincuenta y cinco por ciento de respuestas correctas. Puede usted preguntarse, como yo, qué causa estas diferencias. La capacidad de procesamiento de la memoria de trabajo justifica parte de la variabilidad en la actuación, pero también deben participar otros factores desconocidos. Si usted dice, es una cuestión de inteligencia, tiene razón, porque los test psicológicos de inteligencia correlacionan con la habilidad para el razonamiento. Desgraciadamente, nadie sabe qué subyace en la inteligencia, y muchas de sus medidas dependen de test de razonamiento. Por tanto, su opinión es en cierta medida circular. Un importante reto intelectual para los psicólogos es definir qué subyace en las diferencias en la habilidad para hacer deducciones.

El segundo resultado sólido es que los silogismos varían en su dificultad. Algunos son juegos de niños, como este ejemplo:

Algunos de los autores son panaderos.

Todos los panaderos son jugadores de cricket.

¿Qué se sigue, si se sigue algo?

Casi todo el mundo, desde los niños de siete años a los adultos, extrae la conclusión válida:

Algunos de los autores son jugadores de cricket.

Unos pocos participantes extraen la conclusión válida conversa:

Algunos de los jugadores de cricket son autores.

Pero pocos extraen alguna otra conclusión o responden que no se sigue nada a partir de las premisas. Por el contrario, otros silogismos son tan difíciles que casi nadie obtiene una respuesta correcta para ellos. Si quiere ponerse a prueba, intente este problema:

Ninguno de los artistas es apicultor.

Todos los apicultores son químicos.

¿Qué se sigue, si se sigue algo?

Le daré un momento para pensar en este problema, pero no se sienta abatido si se equivoca. La mayoría lo hacemos. Algunas personas extraen la conclusión:

Ninguno de los artistas es químico.

o su conversa:

Ninguno de los químicos es artista.

Algunas dicen que no se sigue nada determinado sobre los artistas y los químicos. Hay quien extrae la conclusión:

Algunos de los artistas no son químicos.

Pero solo un puñado extrae la conclusión válida:

Algunos de los químicos no son artistas.

Como verá usted más adelante, la teoría de modelos explica esta diferencia de dificultad entre los dos tipos de inferencia.

La teoría postula que una premisa, como por ejemplo:

Algunos de los autores son panaderos.

tiene un único modelo mental que representa un número de sujetos pequeño, pero arbitrario:

autor panadero

autor panadero

autor

...

Cuando los participantes utilizaban para razonar un conjunto de siluetas recortadas, este era el modelo que construían con más frecuencia para esa premisa. En lógica, se considera que el significado de la premisa es que al menos algunos de los autores son panaderos, y por tanto es compatible con que todos los autores sean panaderos. El modelo mental puede ampliarse para representar esta posibilidad añadiéndole más referentes.

Una segunda premisa:

Todos los panaderos son jugadores de cricket.

puede utilizarse para actualizar el modelo anterior. Cada panadero se representa también como un jugador de cricket:

autor panadero jugador de cricket

autor panadero jugador de cricket

autor

...

Este modelo da lugar a la conclusión de que algunos de los autores son jugadores de cricket. Ningún otro modelo para las premisas refuta la conclusión, y por lo tanto es válida.

Un programa que simula esta descripción utiliza procedimientos similares a estos para formar conjunciones de posibilidades (véase el Capítulo 8), porque cada sujeto es similar a un posible modelo. La Tabla 10.3 ilustra cómo funcionan los procedimientos para la inferencia difícil planteada anteriormente. Un principio clave es que dada una premisa que interrelaciona panaderos y jugadores de cricket, pongamos por caso, estando ya los autores y los panaderos representados en un modelo, los sujetos actualizan su representación de los panaderos. De ahí que, en principio, tiendan a no considerar si algunos autores son jugadores de cricket a menos que la relación se establezca por medio de los panaderos. Los modelos de la tabla dan lugar a una única conclusión válida, y la tabla muestra por qué es difícil inferirla. Como revela el programa, los silogismos que dan lugar a conclusiones válidas son de dos tipos: los que dependen de un único modelo mental como el ejemplo anterior y los que dependen de múltiples modelos como el ejemplo de la tabla. Los experimentos han mostrado que quienes razonan son más rápidos y cometen menos errores con inferencias que dependen de un solo modelo que con las que dependen de múltiples modelos. Este resultado se ha replicado en muchos estudios, y ninguno hasta ahora ha dejado de corroborarlo.

Tabla 10.3. Un ejemplo de los procedimientos para el razonamiento silogístico

La primera premisa: Ninguno de los artistas es apicultor, tiene el modelo mental:

artista

artista

apicultor

apicultor

...

con una nota a pie de página según la cual los sujetos representados por la elipsis no son ni artistas ni apicultores.

La segunda premisa: Todos los apicultores son químicos, se utiliza para actualizar este modelo:

artista

artista

apicultor químico

apicultor químico

...

Este modelo da lugar a la conclusión: Ninguno de los artistas es químico, y su conversa.

El procedimiento para buscar modelos alternativos puede añadir nuevas propiedades a un sujeto, dividir un sujeto en dos y reunir dos sujetos en uno. Pero está limitado por las notas a pie de página sobre los posibles tipos de entidad. El procedimiento construye el modelo:

artista químico

artista

apicultor químico

apicultor químico

...

Este modelo refuta las dos conclusiones previas, pero apoya las conclusiones: Algunos de los artistas no son químicos, y su conversa. Este procedimiento construye un tercer modelo:

artista químico

artista químico

apicultor químico

apicultor químico

...

Solo una conclusión se mantiene en los tres modelos: Algunos de los químicos no son artistas.

El primer experimento en el que los participantes sacaron sus propias conclusiones a partir de premisas silogísticas reveló un nuevo fenómeno. Dadas, por ejemplo, las premisas:

Algunos de los autores son panaderos. (A-B)

Todos los panaderos son jugadores de cricket. (B-C)

es más probable que los participantes extraigan la conclusión válida:

Algunos de los autores son jugadores de cricket. (A-C)

que su conversa válida:

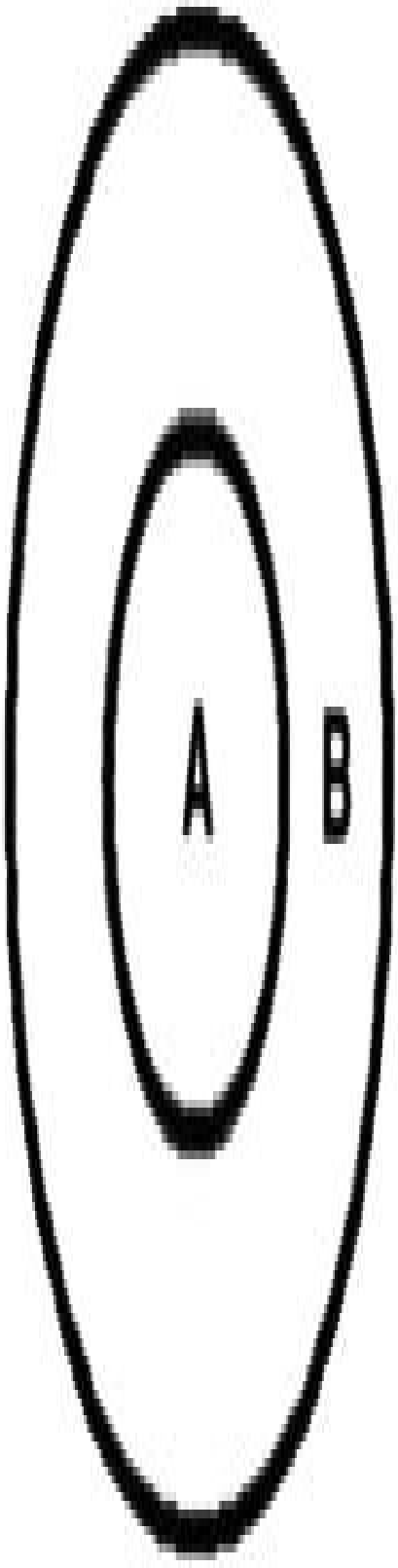
Algunos de los jugadores de cricket son autores. (C-A)

En general, la disposición o “figura” de los términos, como por ejemplo autores, panaderos, y jugadores de cricket, que he simbolizado como A, B y C, sesga las conclusiones que extraen los participantes (véase la Tabla 10.2). Se han adelantado varias explicaciones que compiten en la explicación de este sesgo. La explicación definitiva se debe a Klaus Oberauer. Cuando construimos un modelo para una proposición cuantificada, como por ejemplo: Algunos de los autores son panaderos, representamos en primer lugar un conjunto de autores y luego los actualizamos para asegurar que algunos de ellos son panaderos. Es más difícil hacerlo en el orden opuesto. Por consiguiente, el significado de una oración cuantificada nos lleva a representar en primer lugar su sujeto gramatical. Como puede usted recordar del capítulo anterior, en una premisa sobre relaciones espaciales, como por ejemplo: el triángulo está a la derecha del círculo, ocurre lo contrario. En este caso, preferimos representar en primer lugar el objeto, y buscar a la derecha en la dirección que corresponde al significado de la relación.

* * *

Cuando los participantes inexpertos piensan en voz alta mientras se enfrentan a los silogismos tienen poco que decir. Pero muchos dibujan diagramas conocidos como círculos de Euler. Como ayuda mnemotécnica puede hacer usted el esfuerzo de recordar que los círculos se llaman “de Euler” porque los inventó Leibniz (¡!). Euler se llevó el mérito, porque popularizó el método en sus cartas a una princesa sueca en las que le enseñaba lógica. Los círculos de Euler son similares a los modelos completos en vez de a los modelos mentales. Por

consiguiente, una premisa con la forma Todos los A son B exige dos diagramas independientes, que aparecen en la Figura 10.1. El diagrama de la izquierda representa la posibilidad de que, aunque todos los A son B, algunos B no son A. El diagrama de la derecha, sin embargo, representa que todos los B son A. El uso tradicional de los círculos de Euler exige diagramas que representen todas las posibilidades diferentes para las premisas de los silogismos. Esta exigencia conduce a un gran número de diagramas –dieciséis, por ejemplo, para la fácil inferencia anterior–. No parece que utilicemos los círculos de Euler de esta forma, porque cuando los dibujamos casi siempre dibujamos solamente un único diagrama. Los psicólogos que defienden los círculos de Euler como representaciones mentales han diseñado por tanto formas más eficientes de utilizarlos. Dado que los círculos son invención de un gran matemático, sospecho que los sujetos inexpertos no crean el método por sí mismos, sino que lo recuerdan de sus días escolares. Los círculos de Euler, sin embargo, son una hipótesis legítima sobre la naturaleza de los modelos completos. Los psicólogos no saben si quienes dibujan círculos de Euler, cuando no tienen papel y lápiz, razonan utilizando sus representaciones mentales.


A
B

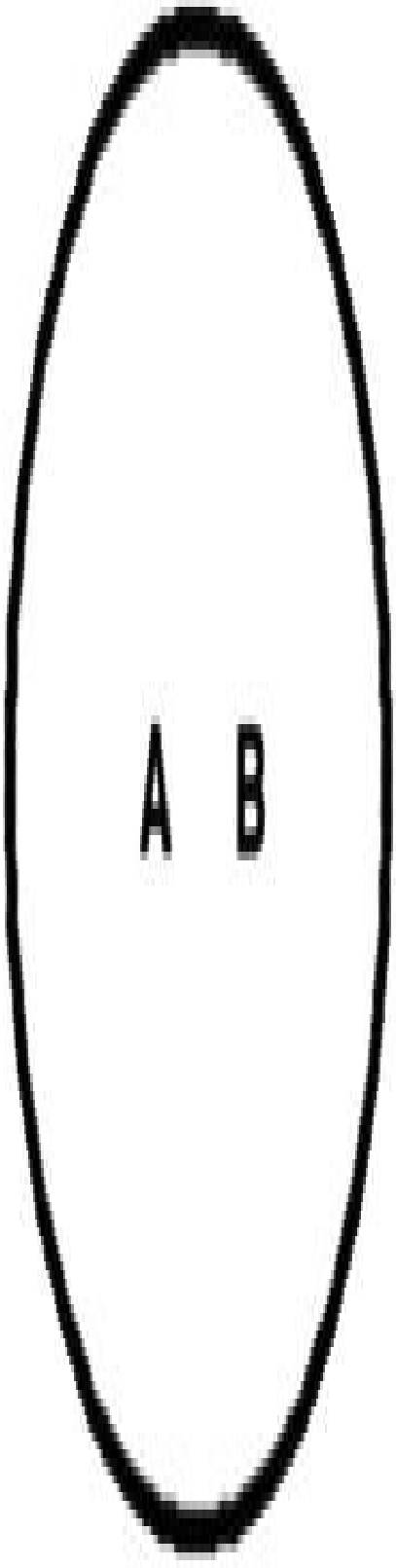

A
B

Fig. 10.1. Los dos diagramas de Euler en forma de círculos que representan una proposición de la forma Todos los A son B.

En el siglo diecinueve, el reverendo John Venn inventó un tipo superior de diagrama para razonar sobre propiedades. La Figura 10.2 presenta un ejemplo. Los diagramas de Venn se confunden a menudo con los círculos de Euler, pero un único diagrama de Venn con tres círculos superpuestos y en el que se utiliza una notación especial para representar posibilidades alternativas es suficiente para cualquier silogismo. Los sujetos inexpertos no dibujan diagramas de Venn, aunque se ha basado en ellos una teoría psicológica.

El razonamiento silogístico válido puede explicarse en términos de modelos mentales, reglas formales, reglas escolásticas como las de la Tabla 10.2, círculos de Euler, diagramas de Venn y aún más principios. Podría suponerse que es imposible averiguar qué sistema utilizamos cuando hacemos inferencias silogísticas. En realidad, las teorías difieren en sus predicciones sobre el proceso de inferencia y sobre la dificultad relativa de los diferentes silogismos. Una tradición que viene de lejos en psicología distingue entre quienes utilizan círculos de Euler y quienes utilizan algún otro método. Pero las pruebas no son decisivas. Los resultados podrían explicarse mediante la teoría de modelos unida a la suposición de que hay quien descubrió los círculos de Euler en el colegio. No existe consenso, y algunos psicólogos ni siquiera creen que razonemos cuando nos enfrentamos a los silogismos. Una salida al impasse es preguntar si hay una teoría que se generalice a otros campos, y si otros resultados podrían proporcionar pruebas cruciales. El siguiente capítulo considera una generalización de la teoría de modelos para inferencias más allá del ámbito de los círculos de Euler o los diagramas de Venn.

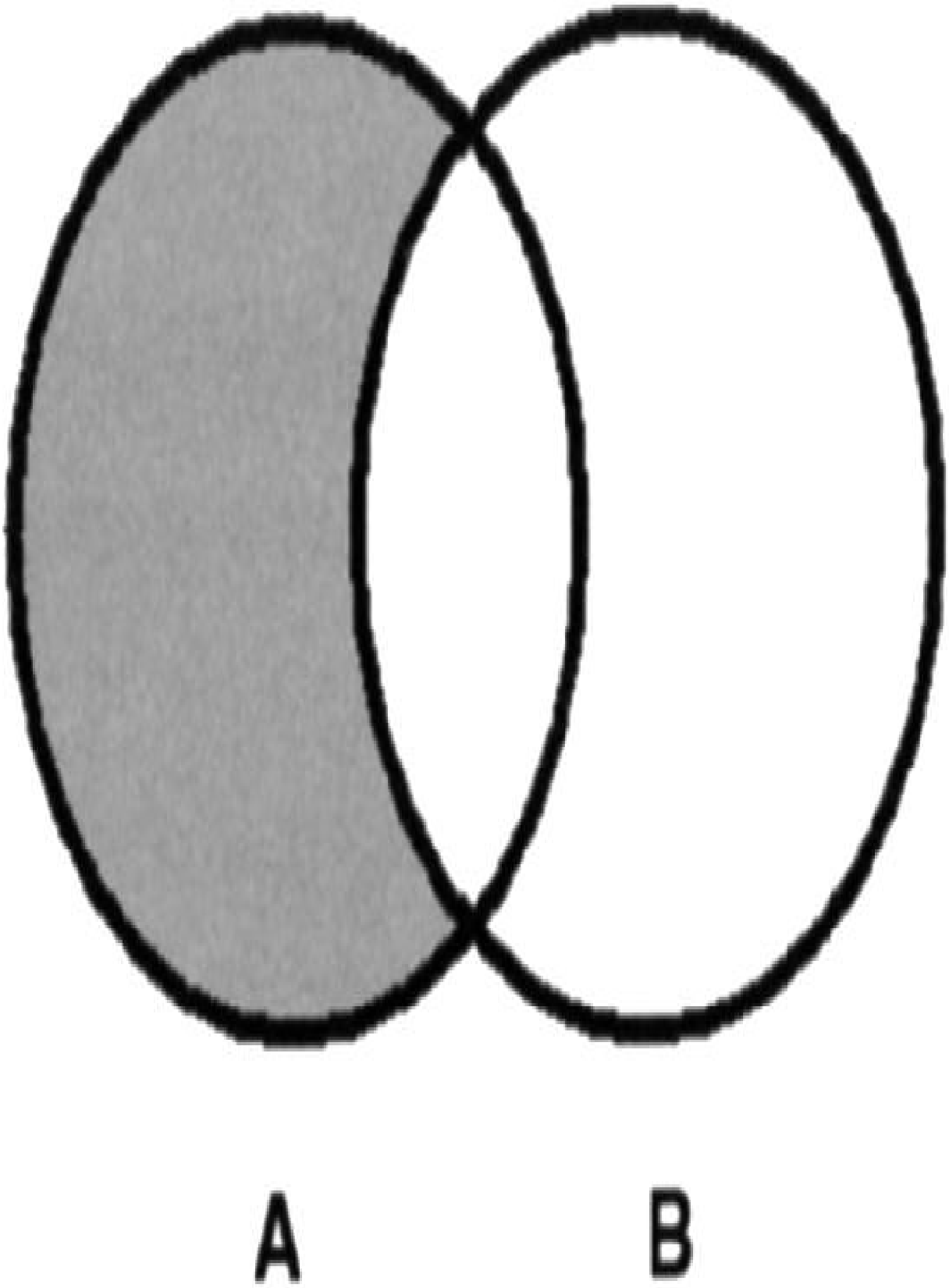

A
B

Fig. 10.2. El diagrama de Venn único que representa una proposición de la forma Todos los A son B. La parte del círculo que representa los A que no son B está sombreada para mostrar que está vacía.

Otras pruebas apoyan la teoría de modelos y van en contra de las versiones alternativas. Si utilizamos modelos mentales, deberíamos sucumbir a inferencias ilusorias sobre propiedades. Estas ilusiones se producen realmente. Supongamos que nos dicen:

Solo una de las siguientes afirmaciones es verdadera:

Al menos algunas de las cuentas de plástico no son rojas, o

Ninguna cuenta de plástico es roja.

¿Es posible que ninguna de las cuentas rojas sea de plástico?

La afirmación del problema dejaba claro, al igual que las instrucciones, que una de las proposiciones es verdadera y la otra proposición es falsa. De acuerdo con el principio de verdad, que definí anteriormente, consideramos la veracidad de cada proposición en sí misma, y no representamos la falsedad de la otra proposición. El modelo de la primera premisa es consistente con la posibilidad de que ninguna de las cuentas rojas sea de plástico, y lo mismo sucede con el modelo de la segunda premisa. De ahí que debamos responder, “Sí”. Aun así, la respuesta es una ilusión. Si la primera premisa es verdadera, entonces la segunda premisa es falsa. Y si la segunda premisa es falsa, al menos algunas de las

cuentas de plástico son rojas, y por tanto algunas de las cuentas rojas son de plástico. A la inversa, si la segunda premisa es verdadera, entonces la primera premisa es falsa, y por tanto todas las cuentas de plástico son rojas. En cualquier caso, la conclusión es imposible. Si le resulta difícil entender esta explicación, ese es el punto más importante. Es difícil pensar en lo que es verdadero y lo que es falso de una vez y al mismo tiempo.

Los problemas de control deberían suscitar la respuesta correcta incluso si quien razona no logra considerar la falsedad. Un ejemplo basado en las mismas premisas anteriores plantea una pregunta diferente:

¿Es posible que al menos algunas de las cuentas rojas sean de plástico?

Debería suscitar la respuesta, “Sí”, porque la apoya el modelo mental de la primera premisa. Esta respuesta es correcta. El estudio examinó varios tipos de problemas ilusorios y problemas de control. Examinó también ilusiones de imposibilidad para las que los modelos mentales apoyan la respuesta errónea, “No”, y problemas de control para los que esta respuesta es correcta. Hay un cincuenta por ciento de probabilidad de adivinar por casualidad la respuesta correcta, y los participantes se comportaron ligeramente mejor que esa probabilidad en los problemas ilusorios (cincuenta y cuatro por ciento correctos). Fueron mucho mejores en los problemas de control (noventa por ciento correctos). La diferencia se mostró en los veinte participantes.

Un estudio posterior valoró un antídoto para las ilusiones. Se les dijo a los participantes que resolvieran el caso en el que la primera premisa era verdadera y la segunda premisa era falsa, y luego resolvieran el caso en el que la segunda premisa era verdadera y la primera premisa era falsa. En realidad, se les dijo que construyeran modelos completos que representaran tanto lo que es verdadero como lo que es falso. El antídoto mejoró la actuación con las ilusiones (del treinta y nueve por ciento al setenta y uno por ciento correcto). Pero el procedimiento exigía un gran esfuerzo y tenía un coste en la realización de los problemas de control, para los que era innecesario. Los resultados en ellos bajaron (del ochenta y siete por ciento correctos al setenta y dos por ciento correctos). Los efectos del antídoto corroboran el principio de verdad. Esos

efectos parecen inexplicables si no razonamos partiendo de lo que resulta verdadero en las posibilidades.

* * *

Las proposiciones cuantificadas se pueden representar en modelos, tanto si el cuantificador es estándar, como por ejemplo “todos” y “algunos”, como si no lo es, como por ejemplo “la mayoría” y “al menos la mitad”. La teoría considera estas proposiciones sobre propiedades como relaciones entre conjuntos de entidades. Si pudiéramos tener presente al mismo tiempo cualquier número de modelos, no nos equivocaríamos. La cruel realidad es que tenemos una fuerte inclinación a razonar en base a un único modelo.

La simple capacidad para extraer una conclusión correcta no discrimina entre las teorías del razonamiento –aunque entre quienes razonan hay algunos tan competentes que la atmósfera de las premisas no puede explicar su actuación–. Un programa de ordenador que implementa la teoría de modelos tiene una propiedad emergente llamativa: para cada silogismo da lugar a una conclusión inicial que se ajusta al cuantificador de al menos una de las premisas. Esta coincidencia muestra que quienes razonan sin el obstáculo de la atmósfera pueden, a pesar de todo, extraer conclusiones que parecen apoyar sus efectos. En resumen, puede que al enfrentarnos a los silogismos la mayoría de nosotros razonemos.

CAPÍTULO 11

¿No somos todos optimistas? El caso del razonamiento complejo

Erase una vez una estación espacial girando en torno a la Luna, y en esa estación espacial había cuatro personas: Anne, Beth, Charles y Diana. Diana quería a Charles, y se cumplía un principio general: todo el mundo quería a cualquiera que quisiera a alguien. Ciertamente, todo el mundo quiere a quien quiere. ¿Quería todo el mundo a Diana? Por supuesto. Ella quería a Charles, y todo el mundo quería a cualquiera que quisiera a alguien. ¿Quería Beth a Anne? Yo diría: no tengo ni idea, e imagino que usted también. Veremos.

Las inferencias de este tipo dependen de sujetos, tales como Diana y Charles, y en relaciones, como "quería". Pero dependen también de una relación entre los miembros de varios conjuntos de sujetos: todo el mundo quería a cualquiera que quisiera a alguien. Todo el mundo, cualquiera y alguien son todos ellos cuantificadores. Cuando cada argumento de una relación es una frase que contiene un cuantificador, la lógica es mucho más complicada que la lógica de las propiedades. Se la conoce como el cálculo "cuantificativo" o "de predicados". Incluye también la lógica de las propiedades y de las conectivas. Es lo suficientemente potente como para demostrar la mayor parte de las matemáticas. Si los cuantificadores, como por ejemplo todo el mundo, se refieren a entidades individuales, como en "a todo el mundo le gustan algunas películas", una inferencia válida siempre se puede comprobar en un número finito de pasos, pero puede ser imposible probar que una inferencia no es válida. Por consiguiente, mientras damos vueltas intentando demostrar una inferencia, no tenemos forma de saber si llegaremos alguna vez a una decisión sobre si es válida o no.

Quizá, debido a su complejidad, los psicólogos han estudiado en pocas ocasiones las relaciones entre conjuntos de sujetos. Los psicólogos que piensan que nuestro razonamiento es formal han sugerido que nuestra mente contiene una versión de la lógica cuantificativa. Vimos en el capítulo anterior que la descripción que hacen de ella solo se enfrenta a los cuantificadores estándar, y hace difícil probar

algunas inferencias triviales, mientras es fácil hacerlo a partir de modelos que representan conjuntos. Mi propósito ahora es explicar cómo podemos utilizar los modelos para hacer inferencias sobre relaciones entre conjuntos.

* * *

La tarea para quien razona es hacer una representación parsimoniosa de múltiples posibilidades. Supongamos, por ejemplo, que yo digo:

Todos los chicos están en el mismo sitio que algunas de las chicas.

¿Quiero decir que los chicos están todos juntos en un solo sitio con algunas de las chicas, o que todos los chicos están repartidos en varios sitios que comparten con las chicas? Probablemente no me importa mucho, y a usted tampoco. Vamos a acordar una interpretación según la cual cada chico está con varias chicas, y puede que no haya un único lugar donde estén juntos. Por el contrario, en la oración:

Algunas de las chicas están en el mismo sitio que todos los chicos

parece tener la interpretación más sólida según la cual algunas de las chicas están en un solo sitio con todos los chicos. Podría usted suponer que yo me preocupo –o al menos los lógicos se preocupan– de buscarle tres pies al gato lógico. Quizá. Pero las diferencias sí tienen consecuencias lógicas. Por ejemplo, si a las oraciones se les dan sus interpretaciones principales, entonces esta inferencia es válida:

Algunas de las chicas están en el mismo sitio que todos los chicos.

Por tanto, todos los chicos están en el mismo sitio que algunas de las chicas.

Si la primera proposición es verdadera entonces lo es la segunda, pero la conversa no se da necesariamente.

Podemos construir modelos mentales para estas proposiciones, y escribí un programa de ordenador que también los utiliza para hacer inferencias. Cada cuantificador exige que, para construir un modelo, se utilice un bucle en el que se repiten operaciones y se incluya un bucle dentro de otro. La Tabla 11.1 ilustra cómo funciona el programa. El programa también puede construir un modelo del significado principal de la oración:

Todos los chicos están en el mismo sitio que algunas de las chicas.

El programa realiza los dos bucles, pero en el orden contrario. Para la oración de la tabla, el bucle para las chicas incluye al bucle para los chicos. Para esta oración, el bucle para los chicos incluye al bucle para las chicas, y el resultado es un modelo del tipo:

| chico chica chica chica | chico chica chica | chico chico chica chica |

Los lógicos tienen un término rimbombante para los bucles de operaciones; los llaman operaciones “recursivas”. Están en el centro de la computación, y exigen una memoria de trabajo para seguir el rastro del proceso de forma que pueda cambiar de un bucle a otro.

Como muestra la tabla, el modelo de la interpretación principal de la oración:

Algunas de las chicas están en el mismo sitio que todos los chicos

Tabla 11.1. Un ejemplo de los procedimientos para construir modelos de relaciones cuantificadas entre conjuntos

El programa construye un modelo del significado principal de la frase:

Algunas de las chicas están en el mismo sitio que todos los chicos.

Empieza con la frase, “algunas de las chicas”. Construye una representación de una chica, e interpreta “están en el mismo sitio que” construyendo una representación de un sitio como una celda de una disposición espacial, e inserta el referente en ese sitio:

| chica | la primera utilización del bucle para las chicas

Interpreta “todos los chicos” repitiendo cíclicamente un conjunto de instrucciones que añaden otro chico al sitio en el que está la chica. De ese modo, el efecto es actualizar el modelo existente mediante una serie de bucles:

| chica chico | la primera utilización del bucle para los chicos

| chica chico chico | la segunda utilización de ese bucle

| chica chico chico chico | la tercera utilización de ese bucle

El número de repeticiones del bucle es pequeño pero arbitrario, y el bucle podría finalizar en este punto. El cuantificador se refiere al conjunto de chicos completo, y por tanto el proceso añade una nota a pie de página respecto a que el modelo representa el conjunto de chicos completo. La frase, “algunas de las chicas” es plural, y por tanto el programa repite el bucle para las chicas. Añade una representación de otra chica en el mismo sitio:

| chica chica chico chico chico | la segunda utilización del bucle para las chicas

En este punto, se realiza de nuevo el bucle para los chicos. Pero en cada ocasión simplemente comprueba que en el modelo está representado todo el conjunto de chicos, y por tanto no añade ningún referente más. Cada bucle para una frase cuantificada verifica siempre en primer lugar si los correspondientes referentes están ya representados en un modelo. Añade nuevos referentes solo si es necesario. El bucle para las chicas puede finalizar en este punto o puede añadir un nuevo referente. También para ellos construye un número pequeño pero arbitrario.

es el siguiente:

| chica chica chico chico chico |

con una nota a pie de página que indica que el modelo representa todo el conjunto de chicos. Los mismos procedimientos permiten a un modelo actualizarse de acuerdo con otra premisa, como por ejemplo:

Ninguno de los chicos está en el mismo sitio que ninguno de los profesores.

El programa hace un bucle para “ninguno de los chicos” que verifica que en el modelo está representado todo el conjunto de chicos, y para cada verificación hace un bucle para “ninguno de los profesores” añadiendo un referente en un sitio distinto. El resultado es este modelo de ambas premisas:

| chica chica chico chico chico | profesor profesor profesor |

El modelo tiene una nota a pie de página para comunicar que representa al conjunto completo de chicos y profesores. El programa utiliza este modelo par extraer una conclusión que relaciona aquellas entidades que no aparecían en la misma premisa:

Ninguna de las chicas está en el mismo sitio que ninguno de los profesores.

Una búsqueda de un modelo alternativo para las premisas construye un nuevo modelo, añadiendo entidades del conjunto que no está representado íntegramente, p. ej.:

| chica chica chico chico chico | profesor profesor profesor chica |

Este modelo refuta la conclusión, pero este modelo y el anterior apoyan una nueva conclusión:

Algunas de las chicas no están en el mismo sitio que ninguno de los profesores.

Ningún modelo de las premisas refuta esta conclusión, y por tanto es válida.

El procedimiento para construir modelos utiliza números arbitrarios de entidades y de esa forma un modelo puede representar muchas posibilidades del mismo tipo. Utiliza bucles de operaciones repetidas para construir interpretaciones para cada cuantificador, y dentro de un bucle se incluye otro con el fin de enfrentarse a cuantificadores múltiples. En principio, este procedimiento se extiende a relaciones entre cualquier número de conjuntos. También se acomoda a los cuantificadores no ortodoxos, como por ejemplo “más de la mitad”. Por tanto, premisas del tipo:

Más de la mitad de las chicas están en el mismo sitio que todos los chicos.

Todos los chicos están en el mismo sitio que más de la mitad de los profesores.

provocan este tipo de modelo:

| chica chica chico chico chico chico profesor profesor | chica | profesor |

El modelo da lugar a la conclusión válida:

Más de la mitad de las chicas están en el mismo sitio que más de la mitad de los profesores.

Una complicación es que hay muchos tipos de premisas, y muchos tipos de modelos de pares de oraciones con dos cuantificadores cada una. Por tanto, no seguiré con los detalles del procedimiento. El punto importante que hay que entender es que la construcción de modelos para estas premisas exige bucles dentro de bucles, en los que cada bucle es responsable de una frase cuantificada, y el orden de los bucles es crucial.

La teoría da lugar a predicciones evaluables, que los experimentos han corroborado. Como en otros campos, es más fácil para nosotros extraer conclusiones a partir de premisas que necesitan un modelo que a partir de aquellas que necesitan múltiples modelos. Por ejemplo, estas premisas que se utilizaron en un experimento:

Todas las cartas de Avon están en el mismo sitio que todas las cartas de Bury.

Ninguna de las cartas de Bury está en el mismo sitio que ninguna de las cartas de Caton.

da lugar a solo un modelo:

| Avon Avon Avon Bury Bury Bury | Caton Caton Caton |

También tendemos a extraer conclusiones erróneas a partir de premisas con modelos múltiples que son consistentes con uno solo de los posibles modelos. Una vez, en televisión, el distinguido psicólogo Colin Blakemore me pidió que

le propusiera una inferencia para extraer. Le propuse una que empieza con la premisa:

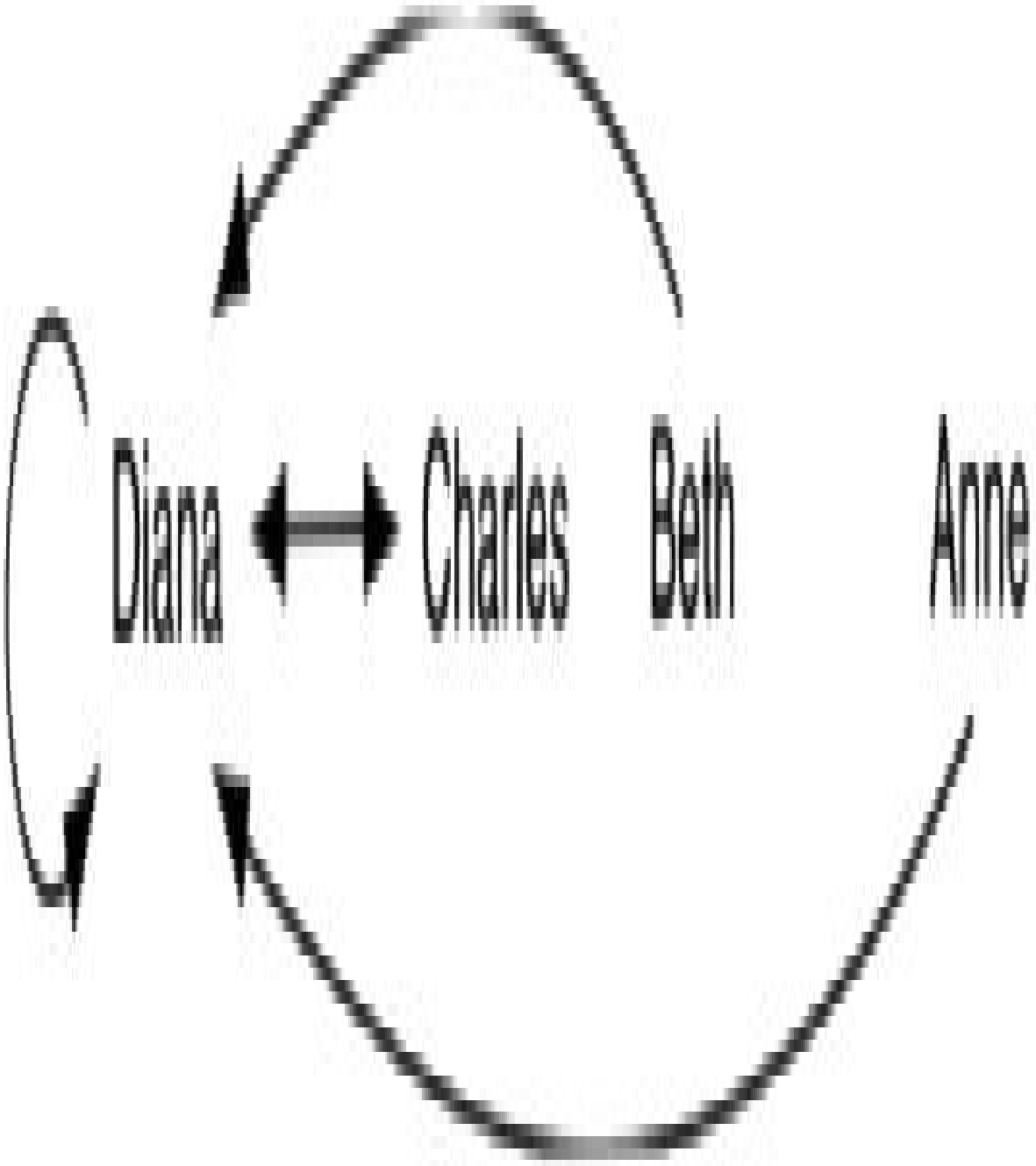
Ninguna de las cartas de Avon está en el mismo sitio que todas las cartas de Bury.

Mientras se repetía a sí mismo la premisa, hizo un gesto con las manos como si estuviera dividiendo las cartas en dos montones separados. Ese gesto, que es visible en la grabación, es un signo claro de alguien que razona construyendo un modelo mental.

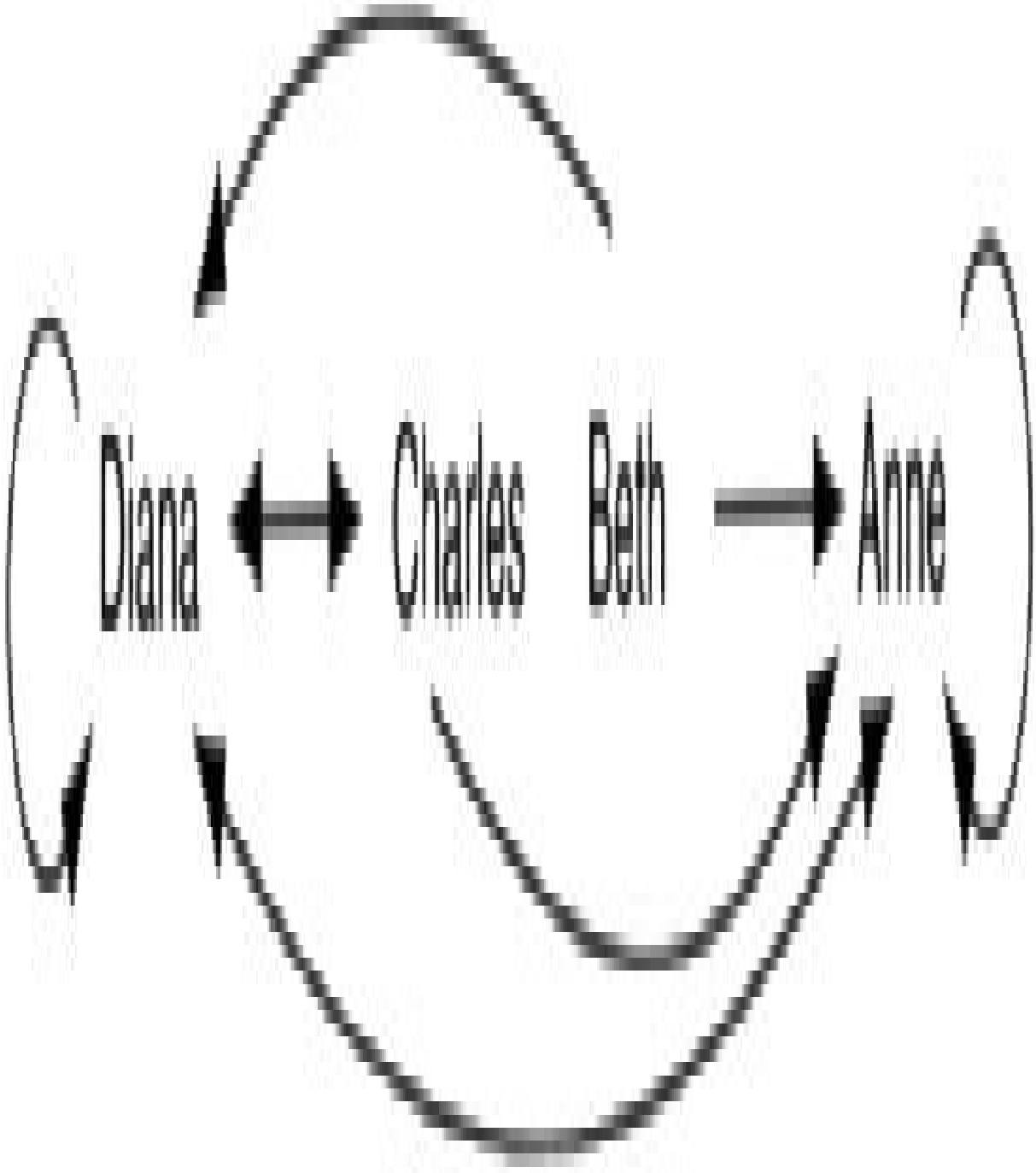
* * *

Dejamos a los cuatro habitantes de la estación espacial sabiendo solamente que Diana quiere a Charles, y que todo el mundo quiere a Diana, porque esta inferencia se sigue de inmediato de la premisa general de que todo el mundo quiere a cualquiera que quiera a alguien. Pero, ¿quiere Beth a Anne? Probablemente su intuición es que esta conclusión no se sigue de las premisas. Las intuiciones son inferencias inconscientes basadas en principios simples –en este caso, que no hay relación entre Beth y Anne, porque no se afirmó ninguna relación entre ellas (ver el Capítulo 5)–. De la misma forma, los participantes en un experimento infirieron que todo el mundo quería a Diana (noventa y ocho por ciento), pero muy pocos (seis por ciento) infirieron que Beth quería a Anne. Sin embargo, los pocos afortunados que infirieron una relación estaban en lo cierto. He aquí por qué. Ya sabemos que todo el mundo quiere a Diana (la inferencia inmediata). Por tanto, Anne quiere a Diana, y por ello Anne quiere a alguien. Podemos utilizar ahora de nuevo la premisa general para inferir que todo el mundo quiere a Anne, y por tanto se sigue a su vez que Beth quiere a Anne. La premisa general tiene la propiedad sutil pero asombrosa de que tan pronto como establecemos que una persona quiere a otra, se sigue que todo el mundo quiere a todo el mundo.

La teoría de modelos predice que las inferencias que dependen de un uso repetido de una premisa –una recursión– deberían ser más difíciles que las inferencias inmediatas. El primer modelo para las dos premisas, incluido el hecho de que Diana quiere a Charles (y a sí misma), es:


Diana
Charles
Beth
Anne

Como muestra este modelo, todo el mundo quiere a Diana. También da lugar a la conclusión de que Anne quiere a Diana. La premisa general se puede utilizar de nuevo para actualizar el modelo:


Diana
Charles
Beth
Anne

Y por tanto se sigue que Beth quiere a Anne. Una tercera utilización de la premisa general da lugar a la conclusión de que todo el mundo quiere a todo el mundo (incluidos ellos mismos), y en este punto el modelo no se puede extender más.

Elimine de su memoria el ejemplo anterior, y considere un nuevo caso que afecta a Alan, Bessie, Carol y David. Usted sabe que Alan no odia a Bessie, y que en este nuevo mundo se cumple esta premisa general:

Todo el mundo odia a cualquiera que odie a alguien.

Dada esta información, ¿se sigue que Carol no odia a David? Nuestra intuición es que no hay información en ningún sentido: podría ser que sí o podría ser que no. Una vez más, sin embargo, nuestra intuición entra en conflicto con una deducción consciente. Supongamos que Carol odiara a David, entonces todo el mundo odiaría a Carol, incluida Bessie. Por tanto, todo el mundo debería odiar a Bessie, y por ello Alan odiaría a Bessie. Pero sabemos que Alan no odia a Bessie, y por tanto se sigue que Carol no odia a David.

Las inferencias que dependen de un uso repetido de una premisa son más difíciles que las inferencias inmediatas. Una razón puede ser que asumimos que una vez que una premisa se ha utilizado en una inferencia, no necesitamos volver a utilizarla. Por consiguiente, tendemos a no expandir los modelos cuando ya hemos construido un modelo que da lugar a una conclusión. El paso que se repite también depende del uso de un modelo inicial, y la representación de este modelo ocupa previamente la memoria de trabajo. Esta carga puede impedir cualquier proceso de inferencia posterior. Sin embargo, en un experimento se les dieron a los participantes conclusiones para evaluar. Escribieron en un papel sus argumentos y pensamientos mientras intentaban entender por qué se seguían las conclusiones a partir de las premisas. En tres cuartas partes de las ocasiones, dieron una explicación correcta de la validez del paso que se repetía, aunque la inferencia negativa que presenté antes era más difícil que la afirmativa. La Tabla 11.2 presenta un protocolo típico en el que el participante, no sin dificultad, finalmente entiende que lo que se sigue es la conclusión afirmativa.

En teoría, podríamos haber razonado sobre estos problemas de una forma totalmente diferente. Antes de considerar la premisa sobre los sujetos, como Charles y Diana, podríamos haber considerado las consecuencias recursivas solo para la premisa cuantificada. Dada la premisa general:

Todo el mundo quiere a cualquiera que quiera a alguien.

podríamos argumentar.

Supongamos que una persona quiere a alguien.

Por tanto, todo el mundo quiere a esa persona.

Por tanto, todo el mundo quiere a alguien.

Por tanto, todo el mundo quiere a todo el mundo.

Por tanto, si una persona quiere a alguien entonces todo el mundo quiere a todo el mundo.

Podríamos entonces combinar esta conclusión con la premisa:

Anne quiere a Beth

para concluir que todo el mundo quiere a todo el mundo. Se sigue inmediatamente que Charles quiere a Diana. Por consiguiente, una posible estrategia es seguir actualizando el modelo de la premisa cuantificada una y otra vez mediante una recursión antes de combinar el modelo con la premisa sobre los sujetos concretos. Ninguna de las justificaciones escritas informó nunca de esta estrategia. Quizá no razonamos de esta forma, porque no nos gusta razonar de manera abstracta. Es decir, preferimos pensar en sujetos concretos, y por tanto nuestro primer paso es interpretar la premisa sobre Charles y Diana. Por el contrario, un alien sobrehumano captaría las consecuencias recursivas de la premisa general. Vería inmediatamente que tan pronto como una persona quiere a otra, todo el mundo quiere a todo el mundo.

* * *

Tabla 11.2. Un protocolo de un participante que hace una inferencia recursiva (traducida del italiano)

El problema era:

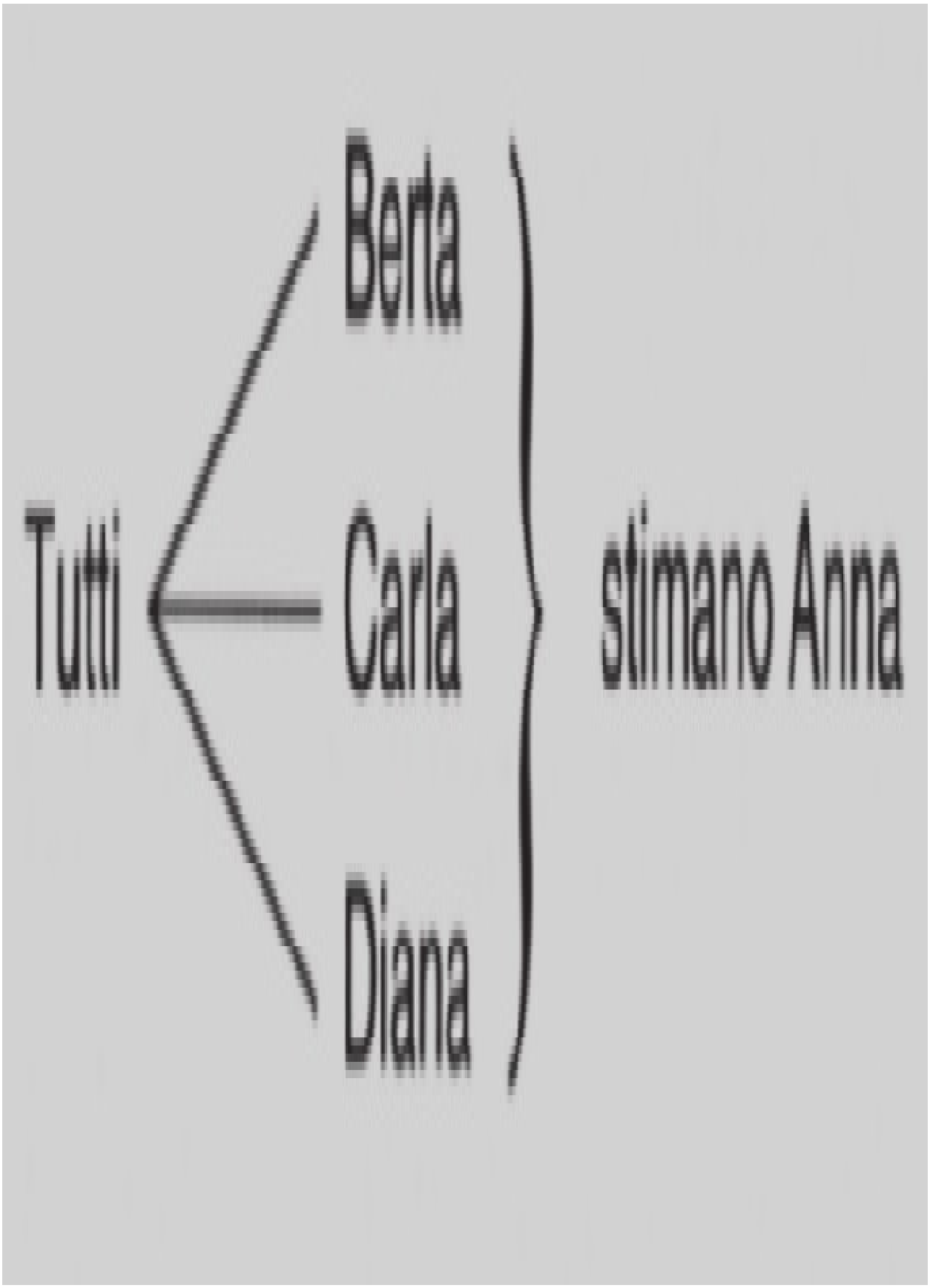
Todos estiman a cualquiera que estima a alguien.

Anna estima a Berta.

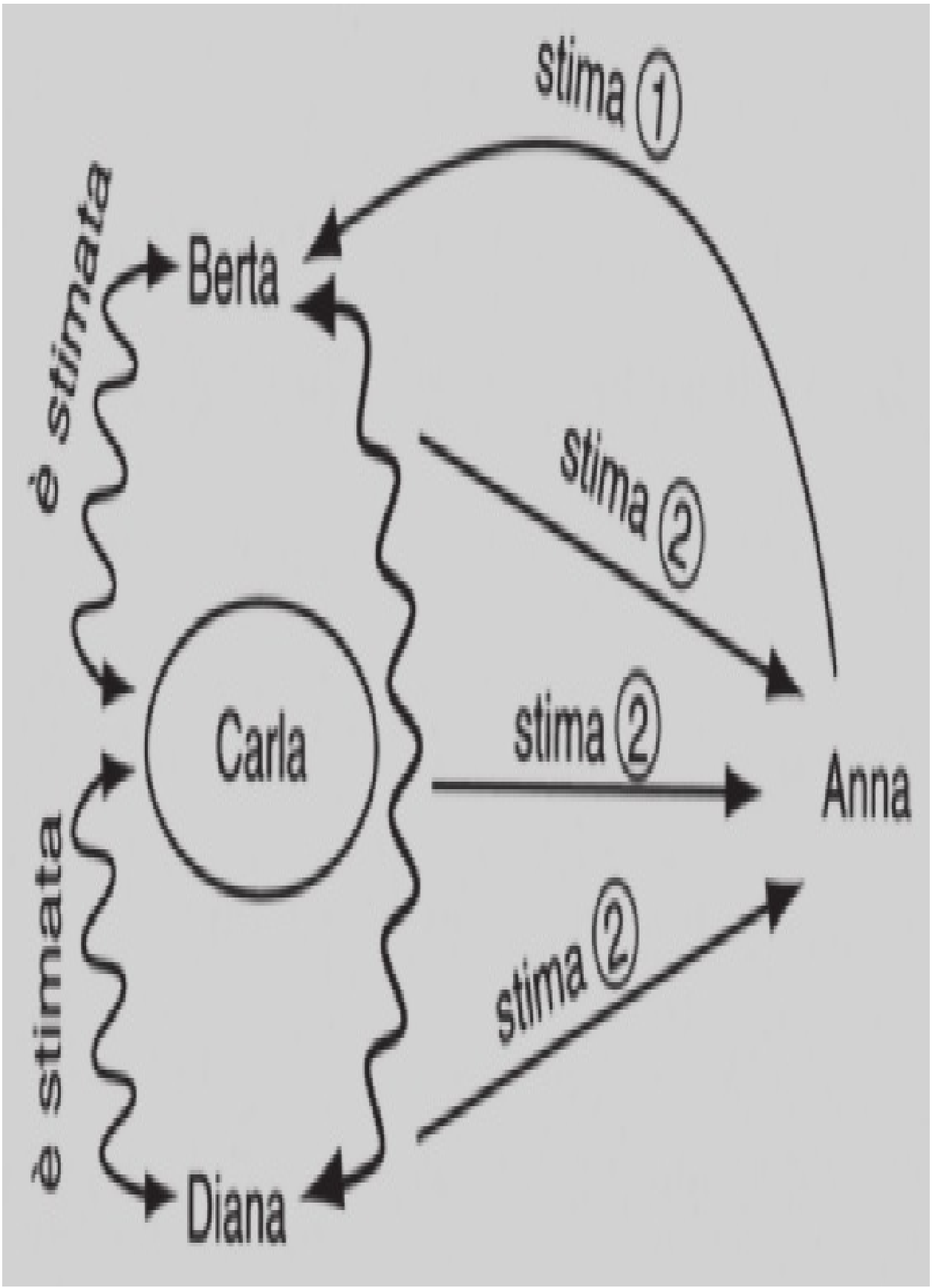
¿Se sigue que Diana estima a Carla?

El participante escribió y dibujó todo lo que aparece a continuación excepto los comentarios entre paréntesis en cursiva:

La conclusión no se sigue necesariamente, porque si asumo la premisa 1, tengo la seguridad de que una persona es estimada si estima a alguien. Por tanto, si supiera que Carla estimaba a alguien podrían aceptar decididamente la conclusión, pero dado que las premisas no dicen nada explícito sobre el hecho de que Carla estime o no estime a alguien la conclusión podría seguirse pero no se sigue necesariamente. A partir de la premisa 2, yo creo que:


Tutti
Berta
Carla
Diana
stimano Anna

(Todos: Berta, Carla, Diana, estiman a Anna).


stima ①
è stimata
Berta
stima ②
Carla
stima ②
Anna
è stimata
stima ②
Diana

Pero no encuentro ninguna relación directa entre Diana y Carla. Entendí la solución después de resolver [otro] problema. Es:

Pero entonces Carla estima a Anna, y por tanto es estimada por todos y también por Diana; por tanto, Diana sí que estima a Carla.

(“stima” = estima; “è stimata” = es estimada. La flecha superior representa la primera premisa; las tres flechas de debajo representan la inferencia inmediata; y las flechas onduladas incluyen la inferencia recursiva.)

Razonamos sobre las relaciones entre miembros de conjuntos de forma similar a como razonamos sobre las relaciones entre sujetos concretos. Construimos modelos, y representamos conjuntos de entidades utilizando conjuntos de referentes. Mantenemos los números reducidos pero arbitrarios, y preferimos trabajar con un modelo simple, en caso contrario las inferencias nos resultan demasiado complicadas de hacer. Las pruebas corroboran esta explicación. Una ventaja de una representación icónica de conjuntos de sujetos es que podemos razonar tanto a partir de cuantificadores del tipo de “más de la mitad” como a partir de cuantificadores del tipo de “todos” y “ninguno”.

¿Se producen las inferencias consideradas en este capítulo en algún sitio que no sea el laboratorio? En realidad, sí que surgen de vez en cuando. Estaba sentado a la salida de una tienda de una cadena comercial, y había un joven sentado a mi lado rellenando un formulario, que yo sabía que era para solicitar un trabajo en la tienda. (Lo había visto en la silla de al lado cuando me senté inicialmente.) Al llegar a un punto, el joven se quedó atascado, y se volvió hacia mí y me preguntó: “¿qué es un optimista?” Yo dije: “Alguien que siempre espera que sucedan cosas buenas.” Estuve a punto de añadir mi definición favorita de optimista, “Un optimista es cualquiera que cree que existen los optimistas”, pero su acento dejaba ver que el inglés no era su lengua materna. Sin embargo, el remate que no pronuncié tiene una interesante propiedad.

¿Es verdad que los optimistas creen que existen los optimistas? Todo aquel al que he planteado esta cuestión piensa durante un momento y luego asiente a la proposición. Como dijo una persona, “Si no pensaras que existen optimistas no serías muy optimista”. Pregunté a una clase de unos cincuenta estudiantes si aceptaban mi definición. Todos lo hicieron. Aun así, no puede ser verdadera, y nuestra incapacidad para entender su fallo sirve como ejemplo del problema que nos supone hacer inferencias que dependen de la recursión –la utilización repetida de una premisa general–. Pregúntese a usted mismo: ¿piensa usted que existen los optimistas? Por supuesto, lo piensa. Yo creo que hay optimistas y usted también. Por tanto, se deduce de mi definición que ambos también somos optimistas. Si otros se dan cuenta de que somos optimistas (porque creemos que existen los optimistas) están en la misma situación que nosotros: piensan que existen los optimistas –a saber, nosotros–, y por tanto ellos también son optimistas. La definición es una infección mental. Extiende el optimismo entre la población como una plaga feliz. Haces circular la definición, y otros sucumben inmediatamente ante ella y se convierten en optimistas. Esta consecuencia de la definición no es obvia. A diferencia de los razonadores sobrehumanos, nosotros no captamos inmediatamente las consecuencias de las repetidas utilizaciones de las proposiciones generales. Preferimos pensar en sujetos concretos, y luego quizá utilizar las proposiciones generales para actualizar nuestros modelos. Este sesgo es conforme con el principio general de la teoría de modelos: los modelos representan sujetos individuales.

¿Por qué es falsa la definición de los optimistas? Los optimistas sí piensan que existen los optimistas, pero también lo piensan muchos pesimistas. Por consiguiente, la definición no se aplica solo a los optimistas. La definición más adecuada de optimista es: una persona que espera que sucedan cosas buenas. Y no todos somos optimistas.

* * *

Muchas oraciones contienen más de dos cuantificadores, y la gramática de nuestro idioma permite un crecimiento indefinido de la complejidad, p. ej.:

Ninguno de los profesores...

Ninguno de los profesores de todos los hijos...

Ninguno de los profesores de todos los hijos de algunos de los empleados...

y así sucesivamente, ad infinitum.

El idioma nos supera. En el momento en que aparecen tres conjuntos cuantificados en una única relación, en la que cada conjunto tiene una infinidad de posibilidades, no existe garantía de que podamos probar que una inferencia no es válida. Y podemos construir rápidamente oraciones sobre las que no podemos razonar correctamente y que no podemos entender. Los psicólogos apenas se han asomado a ese abismo.

Parte IV

Cómo hacemos inducciones

Un modelo abstracto de sucesos

Procede de pasados experimentos

Auden

CAPÍTULO 12

La modulación: un paso hacia la inducción

Utilizamos el significado de las premisas y nuestro conocimiento para construir modelos de lo que es posible, y a veces los ampliamos a modelos completos. Si extraemos deducciones a partir de modelos completos sin cometer errores, entonces nuestro razonamiento es válido. Nuestra capacidad para mantener presentes las cosas es limitada, y por tanto tendemos a razonar con modelos mentales. Desgraciadamente, estos nos pueden llevar por mal camino porque siguen lo que llamé anteriormente el principio de verdad: solo representan lo que es verdadero. A pesar de todo, nuestro método de deducción se esfuerza por ser lógico, y por tanto podría imaginarse que la lógica es la base para la psicología del razonamiento.

Si esa imagen fuera cierta, entonces yo podría terminar este libro ahora mismo. Como quizás haya notado, el libro no termina tan repentinamente, y por tanto puede usted sacar su propia conclusión. En realidad, gran parte de nuestro razonamiento es inductivo y está fuera del ámbito de la lógica. Puede que nuestras conclusiones sean verdaderas, pero incluso con premisas verdaderas no puede existir garantía de su certeza, porque la inducción es falible. Mencioné en el capítulo inicial que los libros de texto definen a menudo la inducción como el razonamiento desde lo particular a lo general. Pero, como señalé entonces, la definición no es totalmente correcta. Podemos hacer inducciones que dan lugar a conclusiones particulares. Y por eso yo defendía esta definición de trabajo: las inducciones van más allá de la información que se da, y descartan más posibilidades que las descartadas por sus premisas. Ese es su sello.

Esta parte del libro tiene como tema la inducción, y voy a defender que siempre depende del conocimiento. En este capítulo daremos un paso hacia la inducción, porque consideraremos cómo el conocimiento modula la interpretación de las oraciones. Un resultado es que conectivas como “o” e “y” dejan de tener sus interpretaciones lógicas, y por tanto nuestras inferencias van más allá de la información literal que se nos da.

Los psicólogos dan a veces la impresión de que la forma lógica es solo una cuestión de gramática. Que, por ejemplo, cualquier oración con dos cláusulas conectadas por “o” es algún tipo de disyunción. los teóricos que defienden una opinión formal tienen a menudo este punto de vista, y en una ocasión el difunto John Macnamara lo representó perfectamente como sigue:

Entendemos por una regla formal, supongo, una regla que se aplica a una cadena en virtud de su forma. Es decir, la regla se puede aplicar siempre que una cadena se describe con una cierta forma... la cuestión de si en la mente de la gente común (sin entrenamiento en lógica) existe una versión psicológica de esta regla se convierte en si tienen una intuición segura, que aplican de la misma forma a cualquier contenido, al que [se aplica la regla]. Supongo que la tienen. Y para mí, este es el final de la cuestión.

Esta visión de la forma como “aplicable a cualquier contenido” está demasiado cerca de la gramática. Por el contrario, se considera que la forma lógica captura la significación de las proposiciones –las condiciones en las que serían verdaderas–. Y los filósofos han sostenido durante más de cien años que la forma lógica de muchas oraciones está bastante alejada de su gramática. Mi propósito inmediato, por tanto, es mostrar por qué una visión gramatical de la forma no es un buen modelo para nuestro razonamiento. las razones son ilustrativas, y conducen a una imagen más exacta de nuestro pensamiento.

* * *

Esta proposición es cierta:

En el colegio, Mo aprendió latín o aprendió griego.

Si descubrimos que no aprendió griego, entonces podemos inferir que aprendió latín. los procesos descritos hasta ahora en este libro dan lugar a esa inferencia. En lógica es válida, y las teorías formales incluyen una regla para hacerla: A o B; No B; por tanto , A. Pero, consideremos esta premisa:

En el colegio, Mo aprendió latín o aprendió algún idioma.

Y ahora supongamos que descubrimos que Mo no aprendió ningún idioma en el colegio. Si aplicamos la regla formal anterior a estas premisas, inferimos: Mo aprendió latín. Pero la conclusión es absurda. Nadie en sus cabales la extraería (excepto un lógico). la razón es obvia. Sabemos que el latín es un idioma: es parte del significado de la palabra. Por tanto, si Mo no aprendió ningún idioma, no pudo haber aprendido latín. Como mínimo, ya no estamos seguros de la intuición de que la regla formal se aplique a cualquier contenido. Este tipo de ejemplo se da tan a menudo que pone en duda la idea de que la forma lógica de la disyunción sea solo cuestión de que aparezcan dos cláusulas conectadas por "o".

Una reacción sensata ante este ejemplo es argumentar que la verdadera forma lógica de la aparente disyunción es:

En el colegio, Mo aprendió algún idioma (o latín u otro idioma).

Esta reacción defiende un método formal de razonamiento, pero, ¿cómo recobramos esta forma lógica? tendríamos que elaborar las posibilidades consistentes con la significación de la proposición. Y una vez que lo hubiéramos hecho, podríamos basarnos también en ellas para razonar. Hemos llegado de nuevo a la teoría de modelos.

¿Cómo elaboramos las posibilidades consistentes con la proposición? la teoría de modelos introduce un principio sencillo para explicar cómo se hace. Nuestro conocimiento del significado de la palabra, "latín", modula nuestra

interpretación de la conectiva. Bloquea la construcción de un modelo para una posibilidad en la que Mo aprendió latín pero no aprendió ningún idioma, porque sabemos que el latín es un idioma. Como resultado, construimos modelos solo para dos posibilidades. En una de ellas, Mo aprendió latín, y en la otra, aprendió otro idioma. Por cualquier camino, aprendió un idioma. la segunda premisa afirma que no aprendió ningún idioma, y por tanto elimina ambas posibilidades. Contrariamente a la regla formal de inferencia, la mayoría de nosotros infiere: Mo no aprendió latín. Este resultado muestra que el significado de una palabra puede modular la interpretación de una conectiva proposicional, y puede bloquear una posibilidad que, en otro caso, el significado lógico de la conectiva permitiría.

He aquí otro problema inferencial:

O Welles no entró en el ascensor o Hearst salió un piso más arriba.

Welles entró en el ascensor en el segundo piso.

¿Qué ocurrió?

Imaginamos una posibilidad en la que Welles entró en el ascensor en el segundo piso, Hearst estaba ya en él, los dos subieron juntos hasta el tercer piso, en consecuencia Hearst salió. Inferimos esta secuencia de sucesos a partir de las premisas y de nuestro conocimiento de los ascensores. Nuestro conocimiento modula la interpretación de la primera premisa, añadiendo información sobre las relaciones entre los sucesos, tanto espaciales como temporales. En una posibilidad Welles entró en el ascensor antes de que Hearst saliera. otra posibilidad es que Welles no entrara en el ascensor. En este caso, no está claro que significaría para Hearst salir un piso más arriba –¿un piso más arriba de dónde?

Un efecto más del conocimiento nos lleva a ampliar nuestros modelos mentales a

modelos completos. Dada la oración:

O en la rueda de la ruleta sale rojo o si no Viv se arruina.

la modulación nos debe llevar a pensar en dos posibilidades completas. En una, en la ruleta sí sale rojo, Viv gana y por tanto no se arruina; en la otra, en la rueda de la ruleta no sale rojo, Viv pierde y se arruina. Debemos representar ambas posibilidades de una forma completa, a diferencia de la interpretación de las disyunciones que no comprometen a nuestro conocimiento, como por ejemplo: "o hay un triángulo o bien hay un círculo". Nuestro conocimiento hace caso omiso del principio de verdad, que describí en el Capítulo 8, y por tanto pensamos sobre ambas cosas, lo que es verdadero y lo que es falso, para las dos posibilidades.

* * *

Tanto el significado como el conocimiento general pueden modular la interpretación de las conectivas proposicionales. Pueden bloquear la construcción de uno de los modelos para una posibilidad. Pueden añadir información a un modelo, como por ejemplo una relación temporal. Y pueden ayudar a generar modelos completos. los dos primeros efectos añaden información: los modelos contienen más información que el significado de las premisas, y esta información procede de nuestro conocimiento. Por tanto, si extraemos una conclusión a partir de los modelos resultantes, estamos yendo más allá de la información contenida en las premisas literales: estamos haciendo una inducción en lugar de una deducción. El tercer efecto permite a las inferencias tener en cuenta información que está presente en las premisas, pero a la que a menudo no se presta atención. Por tanto, ¿cuáles son los procesos que subyacen en la modulación de las premisas? Mi respuesta se ha implementado en un programa de ordenador que simula la modulación.

Nuestro conocimiento a largo plazo se puede representar como modelos de

posibilidades. Por ejemplo, sabemos que el café sabe dulce si se le ha añadido azúcar o un edulcorante artificial. Este conocimiento se puede representar en modelos completos para estas posibilidades: añadiéndole azúcar el café sabe dulce; añadiéndole edulcorante el café sabe dulce; y si no se le añade azúcar ni edulcorante, el café no sabe dulce. tomamos un sorbo de nuestro café matinal y sabe dulce, y por tanto inferimos que tiene azúcar o edulcorante. El proceso en este caso es que un suceso encaja en los modelos de nuestro conocimiento, y da lugar a nuestra conclusión disyuntiva.

En la mayoría de las ocasiones, el conocimiento vigente desde hace tiempo tiene prioridad sobre lo que nos dicen los demás. Sabemos, pongamos por caso, que a Viv no le gusta el café dulce, y por tanto: O el café no está dulce o si no Viv no se lo beberá. Pat nos dice: la taza de café de Viv no está dulce. Pero vemos a alguien añadir azúcar al café. Esta percepción desencadena nuestro conocimiento sobre las consecuencias de esta posibilidad: ahora el café está dulce. Esta proposición contradice el comentario de Pat, pero tiene prioridad sobre el, y da lugar a la inferencia: Viv no se beberá el café.

El proceso que amplía los modelos mentales a modelos completos funciona de una manera análoga. los modelos mentales para el café contiene azúcar o si no edulcorante representan dos posibilidades. Estas posibilidades desencadenan las del conocimiento y por tanto cada modelo se amplía a un modelo completo que representa que en ambas posibilidades el café está dulce.

El proceso que añade información a los modelos, como por ejemplo una relación temporal, es, en principio, fácil de entender. Por ejemplo, tenemos un modelo de los ascensores que se desarrolla en una secuencia cinemática. Sabemos que traslada a la gente de un piso a otro en los edificios. Conocemos el protocolo. los pasajeros presionan un botón para llamar al ascensor a su piso. llega y se abren las puertas. Entran y presionan el botón del piso al que quieren ir a menos que alguien ya lo haya pulsado. las puertas se cierran y el ascensor les lleva allí, parándose quizá por el camino en otros pisos. Sabemos también que en raras ocasiones, como en todas las máquinas, las cosas pueden funcionar mal. Alguien nos dice:

Hearst entró en el ascensor y salió un piso más arriba.

Nuestro conocimiento del significado de “ascensor” desencadena nuestro modelo de cómo funcionan los ascensores. Hacemos una simulación mental, parte por parte, de la secuencia de sucesos: Hearst entra en el ascensor en un piso, la puerta se cierra, el ascensor sube un piso, las puertas se abren y Hearst sale. Inferimos esta secuencia cinemática a partir de nuestro conocimiento, pero es una inducción. Hemos ido más allá de la información presente en la proposición para hacer una secuencia de inferencias falible. Puede ser, por ejemplo, que Hearst abriera las puertas en un piso forzándolas, se saliera al eje del ascensor, subiera por los cables entrando al ascensor a través de una puerta que hubiera en el suelo. El ascensor con las puertas abiertas estaba parado entre dos pisos. Él se puso sobre los hombros de alguien y salió del ascensor un piso más arriba. todos los sucesos de nuestra simulación resultaron ser falsos. No es imposible, pero es muy improbable. Por eso es por lo que compensa utilizar nuestro conocimiento: casi siempre es correcto. Pero es difícil pensar cómo podríamos hacer estas inferencias sin recurrir a un modelo mental de los ascensores.

* * *

La modulación afecta también a los cuantificadores. Una señal habitual en Inglaterra a la entrada de una finca en el campo es:

Quien entre sin permiso será procesado[1].

Sabemos que deja abierta la cuestión de si habrá o no alguien que entre sin permiso. Por el contrario, un folleto de un curso universitario dice:

Los estudiantes graduados serán financiados con becas durante su primer año.

No tenemos dudas de que hay estudiantes graduados en la universidad.

El conocimiento puede determinar también las relaciones entre conjuntos de sujetos. En lógica, una oración como al menos algunos phykods son entalics deja abiertas dos cuestiones: puede ser que todos los phykods sean entalics. Y también puede ser que todos los entalics sean phykods o que solo algunos lo sean. Dado que no sabemos nada de los phycods o los entalics –una asunción segura, dado que me he inventado estos términos, no tenemos forma de conocer la relación concreta entre los dos conjuntos–. Pero en el caso de la oración: algunos animales son perros, sabemos las dos cosas: que algunos animales no son perros y que todos los perros son animales. Por el contrario, con la oración: algunos conservadores son hombres de negocios, sabemos que no todos los hombres de negocios son conservadores. los experimentos han confirmado esta clase de diferencias en la interpretación de proposiciones cuantificadas. Con "muchos" y "unos cuantos" se producen efectos similares. Nuestro conocimiento del tamaño de un conjunto tiende a afectar a nuestra estimación respecto al número de sujetos a los que nos referimos en proposiciones como por ejemplo:

A muchos franceses les gusta el vino. (Estamos hablando de millones.)

Muchos senadores aceptan regalos de los miembros de los lobbys. (Estamos hablando de decenas.)

En lógica, "cada" se refiere a todas las cosas del ámbito que se discute. Por tanto, una proposición como por ejemplo:

La abuela odia a cada miembro de su familia

significa que la abuela también se odia a sí misma porque ella pertenece a su familia. Pero, el conocimiento bloquea con frecuencia estas interpretaciones. los participantes en los experimentos a menudo las pasan por alto –hasta el punto de que si en un estudio se necesita que se tengan en cuenta, el experimentador tiene que explicar su existencia a los participantes–. la moraleja es que el conocimiento modula la interpretación de los cuantificadores. Estos tienen también significados lógicos, pero el conocimiento puede bloquear una determinada posibilidad o, por el contrario, hacerla más destacada.

* * *

La modulación se ha demostrado en estudios en los que los participantes enumeran lo que es posible a partir de una serie de oraciones. Estos estudios mostraron que oraciones con la misma gramática pero con contenidos diferentes realmente dan lugar a los efectos de modulación esperados. oraciones como: Pat está en Río o está en Noruega, provocan el conocimiento de que una persona no puede estar en dos sitios al mismo tiempo, y por tanto el que una cláusula sea verdadera implica que la otra cláusula sea falsa. Por consiguiente los participantes son más rápidos y extraen más conclusiones válidas de la forma:

Pat está en Río o está en Noruega.

Por tanto, si Pat no está en Noruega entonces está en Río.

Sin embargo, se frustran con inferencias, como por ejemplo:

Pat está en Río o está en Brasil.

Por tanto, si Pat no está en Brasil entonces está en Río.

En contra de la forma disyuntiva de la premisa, el que la primera cláusula sea cierta implica que lo sea la segunda.

* * *

¿Por qué no es lógica la psicología? Para responder a esta pregunta, quiero reunir temas que ya ha encontrado usted en este libro y enfrentarme a los resultados presentados en este capítulo. El cálculo cuantificativo, que presenté en el capítulo anterior, captura las implicaciones entre oraciones (en un lenguaje formal), mientras que el razonamiento es un proceso de extracción de conclusiones a partir de premisas. El cálculo nos dice poco más sobre el razonamiento de lo que la aritmética nos dice sobre contabilidad. Permite un número infinito de conclusiones válidas a partir de cualquier conjunto de premisas, incluyendo muchas conclusiones estúpidas, tales como una mera conjunción de las premisas. No tiene nada que decir respecto a las conclusiones concretas que tienden a extraer los humanos que razonan o sobre su tendencia a decir que no se sigue nada a partir de ciertas premisas. Por tanto, los psicólogos que suponen que la mente contiene una versión tácita del cálculo proporcionan una “máquina” inferencial que utiliza reglas del cálculo para obtener conclusiones. también deben idear formas de evitar derivaciones que den lugar al tipo de inferencias válidas pero triviales que no haríamos nunca.

Cuando en la vida real creamos un argumento, es difícil analizarlo lógicamente. Muchas inferencias están más allá del ámbito del cálculo cuantificativo clásico, porque se refieren a conclusiones sobre posibilidades. Y damos por hechos los significados de las palabras, su contexto y el conocimiento que tenemos acumulado. Por ejemplo, sabemos qué significa “más alto que”, y por eso podemos utilizarlo para construir un modelo a partir del cual puede o no surgir una conclusión transitiva. la transitividad no se puede capturar en un postulado del significado para “más alto que”, porque depende de la proposición como un

todo. Por ejemplo, un cambio en el tiempo verbal de una premisa a otra puede violar la transitividad, como en estas dos premisas (del Capítulo 9):

Cate es más alta que Belle.

Belle era más alta que Alice (hasta que Alice bebió la poción mágica).

Suponemos una base común de conocimiento y creencias. Por ejemplo, una vez pedí a alguien que diera argumentos en favor de que el gobierno local subvencionara la ópera. Respondió que el arte es bueno para la vida cultural y el comercio. lógicamente, en el argumento han desaparecido varios enlaces, que puedo tratar de detallar:

El arte es bueno para la vida cultural y el comercio.

Cualquier cosa que es buena para la vida cultural y el comercio debería subvencionarse.

La ópera es un arte.

Por tanto, la ópera es buena para la vida cultural y el comercio.

Por tanto, la ópera debería subvencionarse.

Exponer de esta forma un argumento, como una prueba formal, es un poco ridículo, pero la lógica lo exige. la primera premisa es lo que afirmó el hablante. la segunda premisa es quizá una de sus creencias generales, que da por hechas. la tercera premisa se basa en el conocimiento del significado de la palabra, “ópera”. De ahí que los argumentos cotidianos dependan de los significados de las palabras, el conocimiento general y las creencias. la consiguiente dificultad para hacer análisis lógicos la confirma el hecho de que no exista ningún programa de ordenador que lleve a cabo esa tarea.

Los ejemplos de este capítulo son más realistas que las oraciones artificiales de los experimentos. lo que los hace más naturales es que sus contenidos son proposiciones sensatas sobre asuntos cotidianos. Sin embargo, un corolario es que el significado de sus cláusulas individuales y el conocimiento general sobre las situaciones a las que se refieren puede modular la interpretación de las conectivas. Unos usos se corresponden con los significados de los términos en lógica y otros no. la potencial modulación de las conectivas y los cuantificadores significa que nuestro sistema para interpretar las oraciones debe estar siempre a la caza de estos efectos. Debe examinar siempre los contenidos de las oraciones para comprobar si ellos y el conocimiento que provocan modulan la interpretación. Este paso se produce incluso para oraciones que resultan tener una interpretación lógica. El sistema para interpretar conectivas proposicionales no puede trabajar en la forma en que lo hace la lógica, teniendo en cuenta solamente la verdad y la falsedad. El sistema trabaja con modelos de posibilidades, no con valores de verdad, y debe tener en cuenta el contenido y el conocimiento. Por eso es por lo que el proceso de interpretación no es nunca pura lógica. El hecho de que la modulación pueda añadir relaciones temporales y espaciales entre los sucesos descritos en una oración significa que las conectivas proposicionales tienen un número indeterminado de interpretaciones diferentes.

Una reacción de algunos teóricos ante las discrepancias entre lógica y razonamiento es sostener que la lógica no tiene nada que ver con las inferencias de la vida. Estos teóricos están matando moscas a cañonazos. la validez es importante en la vida cotidiana, y es un concepto lógico. otra reacción es salirse de la lógica clásica e invocar otro tipo de lógica, incluidas las lógicas que permiten que se retiren las conclusiones, que discutiré en un capítulo posterior.

Podría usted suponer que al menos en las habilidades mentales de los lógicos y los matemáticos sí que subyace la lógica. Algunos de ellos también lo han pensado, pero esas disciplinas exigen algo más que la sola deducción. El genio

indio Srinivasa Ramanujan, que murió en 1920, dejó tras de sí un cuerpo de trabajo que continúa ocupando a los matemáticos, pero la pruebas no fueron cruciales para sus muchas conjeturas. Como escribió sobre él otro gran matemático, G. H. Hardy: “[él] combinaba un poder de generalización, una sensibilidad para la forma, una capacidad para la rápida modificación de sus hipótesis, que eran a menudo realmente sorprendentes, y hacían que, en su propio campo concreto, no tuviera rival en su tiempo”.

Todos nosotros tenemos algún poder de generalización. Y una razón final para rechazar la lógica como modelo de nuestro razonamiento es el que tantas de nuestras inferencias no sean deductivas, sino inductivas. Por ejemplo, cuando detectamos un conflicto entre los hechos y nuestras creencias, tratamos de razonar nuestra vuelta a la consistencia. Intentamos modificar nuestras creencias para hacerlas consistentes con los hechos y crear un diagnóstico convincente respecto a qué era erróneo. la lógica no puede decirnos qué creencia abandonar, o cómo resolver el conflicto. la forma en que lo hacemos es un asunto al que volveré más adelante, e ilustra la importancia del razonamiento que no es deductivo.

* * *

La modulación es el puente desde la deducción a la inducción. En la vida, cruzamos el puente sin esfuerzo o sin ser conscientes. lo que nos resulta más difícil es evitar que nuestro conocimiento influya en el razonamiento, e intentar hacer deducciones estrictas. Normalmente, los significados de las palabras, las entidades a que se refieren y nuestro conocimiento del mundo modulan, todos ellos, nuestra interpretación de las conectivas y los cuantificadores. Una consecuencia es que el proceso de interpretación nunca puede desarrollarse de una forma lógica idealizada. otra consecuencia es que la modulación afecta al proceso de inferencia: posibilidades implícitas que rara vez nos importa tener en cuenta pueden convertirse en importantes como resultado de nuestro conocimiento. la inducción depende del conocimiento, y la omnipresencia de la introducción del conocimiento en los modelos significa que muchas de las inferencias de la vida, si no la mayoría, son inductivas.

Notas al pie

[1] Nota de la trad.: la señal correspondería al aviso: Prohibido el paso, propiedad privada, pero se ha conservado la traducción literal para mantener el sentido del comentario.

CAPÍTULO 13

El conocimiento y las inducciones

En mi primer viaje a Italia hice algo que nunca había hecho antes. Entré en un bar italiano a tomar un café. Me quedé entre el gentío de la barra hasta que conseguí atraer la atención de uno de los hombres que hacían los cafés, y grité en una aproximación al italiano: “un capuccino per favore”. Él me hizo señas de que me fuera. lo intenté de nuevo: me volvió a hacer señas de que me fuera. Y entonces se volvió hacia el cliente que estaba de pie a mi lado y le sirvió. Yo estaba haciendo algo mal. Pero, ¿qué? Me di cuenta de que el hombre que estaba detrás de mí tenía en la mano un trocito de papel. Ciertamente, las personas que esperaban ser atendidas tenían esos trocitos de papel, que agitaban ante los camareros. Parecían ser recibos de una caja registradora, e inferí que todos los tenían. Sin embargo, nadie pagaba a los camareros. Se deducía que debían haber pagado a alguna otra persona. Miré a mi alrededor, y más allá, en la dirección en que el camarero me había mandado, había una mujer sentada ante una caja registradora y una fila de gente esperando para pagarle. Eso era. Primero se pagaba y se llevaba el recibo a la barra para encargar la consumición. Este sistema sigue utilizándose en muchos bares de Italia, pero yo no lo conocía hasta ese momento.

Mis pensamientos en el bar me llevaron menos tiempo del que le ha llevado a usted leer cómo los he descrito, pero ilustran lo que quiero discutir. lo fundamental en ellos es una serie de inducciones. Recuerda usted que si las premisas son verdaderas, una deducción da lugar a una conclusión que debe ser verdadera, mientras que eso no sucede en el caso de una inducción –la conclusión puede ser verdadera, pero no existe garantía–. las inducciones resolvieron mi problema respecto a cómo conseguir un café. En Inglaterra si un camarero hubiera rehusado servirme una pinta de cerveza en un pub, me habría ido enfurruñado. Pero, en el momento en que fui al bar italiano, ya me había formado una impresión de que los italianos eran simpáticos y serviciales con los extranjeros. Esta impresión también era una inducción. Había experimentado varios actos de amabilidad en los que las personas habían sido simpáticas y

serviciales, y por tanto había inferido, a partir de esta pequeña muestra, que los italianos en general tenían la misma disposición. No podría haber estado de acuerdo con la proposición de que todos los italianos son simpáticos y serviciales. Pero era una asunción que daría por cierta hasta que comprobara que era falsa, y por tanto no inferí que el camarero estaba siendo hostil hacia alguien claramente extranjero.

Hice la inducción de que en los bares italianos uno debe pagar primero y pedir después. Y esta generalización fue el resultado de una única experiencia. Era una generalización porque se aplicaba a un número indeterminado de bares italianos. Era una inducción, puesto que podía haber sido falsa, es decir, mi conclusión no era una deducción a partir de las premisas. las premisas –o sea, el bar– podría haber sido excepcional, y cualquier otro lugar podría haber funcionado con las mismas normas de los bares de Inglaterra, donde el pago llega al final, si uno sobrevive al café. Una razón por la que dudé que el bar italiano fuera excepcional era que si lo hubiera sido, otros clientes habrían estado tan descolocados como lo estaba yo, y habrían tenido que dar el mismo rodeo –a la barra, a la caja registradora y de vuelta a la barra–. Eso no sucedía. la clientela actuaba de una forma correcta –entraban, pagaban en la caja registradora e iban a la barra con sus recibos a pedir sus cafés–. Por tanto, hice una deducción a partir de una premisa verosímil: Si el bar fuera excepcional entonces otras personas estarían descolocadas. Pero, como podía ver, los demás no estaban descolocados; por tanto, el bar no era excepcional. Y esta deducción apoyaba mi inducción.

otra razón a favor era que inferí una explicación para el modus operandi del bar. El lugar estaba muy lleno, y en la barra solo había dos hombres trabajando. Estaban manejando las máquinas de café sin pausa y no tenían tiempo de hacer las notas o dar el cambio en una caja registradora. Una solución eficiente era, por tanto, que los clientes presentaran el recibo, mostrando qué habían pagado, y que los camareros prepararan las bebidas correspondientes.

Cuando nuestro razonamiento va más allá de las premisas hasta aumentar la cantidad de información, estamos haciendo una inducción. Mi generalización, según la cual los bares en Italia exigían a los clientes pagar antes, era una inducción. Mi inferencia de una explicación para este modo de funcionamiento también es una inducción, y algunos teóricos se refieren a las inducciones de explicaciones como abducciones. Y cuando hemos hecho una abducción para explicar una inducción, como yo hice, nuestra creencia en la inducción tiende a aumentar. No solo hemos hecho una generalización, sino que también la hemos

explicado. Y esta es una buena razón para creerla. Algunos filósofos han sostenido que siempre que hagamos cambios en nuestras creencias deberían ser tan pequeños como fuera posible. Dudo que sigamos siempre este consejo o incluso que debamos seguirlo. Nuestra propensión a hacer inducciones y a explicarlas nos lleva a menudo a hacer cambios bastante grandes en nuestras creencias.

En el caso del bar existe una observación psicológica sorprendente, y creo que universal. Cuando razonamos, no somos conscientes de estar saltando de inducciones a deducciones, a abducciones y otra vez hacia atrás. Por el contrario cuando pasamos, pongamos por caso, de leer un soneto a tratar de escribirlo, o de seguir una demostración matemática a tratar de crearla, somos conscientes de un cambio. El razonamiento parece ser de una pieza. Quiero explicar cómo funciona la inducción. Y por tanto voy a describir qué se considera como una inducción y a defender que la abducción es un caso especial de inducción. Me centraré luego en aquellas inducciones “puras” que dan lugar a generalizaciones y propondré una teoría respecto a cómo las hacemos. En el capítulo siguiente explicaré las abducciones.

* * *

Cuando hacemos inducciones nuestro propósito es alcanzar conclusiones nuevas y parsimoniosas que al menos sean verosímiles dadas nuestras premisas y nuestro conocimiento. Pero en la vida real nuestro razonamiento no se produce en un vacío intelectual. Razonamos al servicio de nuestros objetivos. Queremos un capuchino y razonamos para conseguirlo. Nuestros objetivos restringen los tipos de conclusiones que tratamos de inferir. las mismas premisas con diferentes objetivos nos llevan a conclusiones diferentes.

Un tipo normal de inducción de la vida real da lugar a una conclusión sobre un suceso particular. Por ejemplo, cuando el ferry de coches The Herald of Free Enterprise partió del puerto belga de Zeebrudge el 6 de marzo de 1987, el capitán del navío indujo de manera verosímil que las puertas de proa habían sido cerradas. Siempre habían sido cerradas hasta entonces, y no había pruebas de lo contrario. El oficial jefe hizo la misma inducción, lo mismo hizo el

contramaestre. Pero el ayudante del contramaestre, cuyo trabajo era cerrar las puertas, estaba profundamente dormido en su litera. Y no había cerrado las puertas. En cuanto el ferry dejó el puerto, el mar lo invadió a través de las puertas abiertas. El barco volcó y se hundió, y 188 personas se ahogaron. la inducción puede suponer un riesgo terrible.

Otro tipo de inducción conduce a una generalización, como por ejemplo mi inferencia de que en los bares italianos se paga antes de pedir. las leyes científicas son también generalizaciones inductivas. la primera ley de Kepler sobre el movimiento de los planetas, por ejemplo, establece que un planeta se mueve en una órbita elíptica alrededor del Sol, estando el Sol situado en uno de los focos de la elipse. las teorías científicas van más allá de estas regularidades a fin de explicarlas en términos más básicos: la teoría de la gravitación de Einstein explica las órbitas planetarias en relación con los efectos de la masa de una estrella sobre la curvatura del espacio-tiempo. El difunto Sir Karl Popper, un filósofo de la ciencia, sostenía que la inducción no desempeña ningún papel en el pensamiento científico. la ciencia, sostenía, se basa en conjeturas que están abiertas a la falsificación, y los orígenes de estas conjeturas no son de la incumbencia de los filósofos. En realidad, en contra de lo que dice Popper, la historia de la ciencia contiene muchos casos en los que los científicos hicieron observaciones sistemáticas, y buscaron generalizaciones parsimoniosas de sus resultados (también conocidas como leyes científicas). Por ejemplo, la ley de Marbe en psicología –por si usted se pregunta si hay alguna ley psicológica– establece que cuanto más habitual es una asociación con una palabra dada más rápidamente tiende a hacerse. Por ejemplo, “blanco” es la respuesta más frecuente, y más rápida, a la palabra “negro”. Marbe indujo la ley a partir de observaciones sobre respuestas asociativas.

La abducción de una explicación también aumenta la información, y por tanto la abducción es un tipo de inducción. ¿Cómo se puede explicar la diferencia entre abducciones que dan lugar a explicaciones e inducciones que no lo hacen?, podría preguntarse usted. la clave es el producto final. la inducción que hice a partir del bar para los bares en general no introdujo ningún concepto nuevo además de los que se referían a conseguir un café. Reordenó la secuencia temporal para pedir, consumir y pagar. Mi explicación, sin embargo, se refería a la carga de trabajo de los camareros y a su falta de tiempo para manejar el dinero. Estos eran conceptos nuevos más allá de aquellos que se referían a cómo conseguir un café y pagarlo.

La relación entre una explicación satisfactoria y la generalización que explica es mucho más problemática. Una opinión es que la generalización es una consecuencia lógica de la explicación: podemos deducir la generalización a partir de la explicación. Varios filósofos han cuestionado esta opinión, y puede ser demasiado contundente. Pasa por alto el núcleo de la psicología. Podemos describir un suceso sin entenderlo, pero no podemos explicar un suceso a menos que presuntamente lo comprendamos en alguna medida. las descripciones nos permiten, como máximo, remedar mentalmente un objeto, mientras que las explicaciones nos permiten construir un modelo mental del mecanismo que subyace en él. Podemos considerar el suceso por partes: sabemos qué lo causa, qué resultados produce, cómo influir en él, controlarlo, iniciarlo o prevenirlo, cómo se relaciona con otros sucesos o cómo se les parece, cómo predecir su inicio y desarrollo, cuál es su estructura interna o subyacente, cómo diagnosticar sucesos no habituales, y, en ciencia, cómo relacionar el campo como un todo con otros. las explicaciones científicas pueden hacer uso de nociones teóricas que son inobservables, o que están en un nivel físico inferior a las descripciones de los sucesos. Una explicación da cuenta de lo que no entendemos en función de lo que entendemos: no podemos construir un modelo explicativo si no disponemos de los conceptos clave.

La inducción de una generalización está a menudo tan bien descrita como la inducción de un concepto. En el ejemplo anterior, yo (y tal vez usted también) adquirí un conocimiento sobre el concepto:

Los bares italianos que tienen cajeros.

Estos conceptos ad hoc se construyen de una manera inductiva a partir de componentes más básicos, como los conceptos de bares y cajeros. Seguimos adquiriendo conceptos similares a lo largo de toda nuestra vida. Algunos los aprendemos a partir de la experiencia directa y otros a partir de descripciones. No podemos adquirir el concepto completo de un color, una pieza de música, o una emoción sin una relación directa con ellos. Pero podemos aprender sobre quarks, genes y el inconsciente a partir de sus descripciones.

En resumen, las inducciones pueden referirse a un suceso concreto o a temas en

general. Pueden ser descripciones o explicaciones. Incluyen la adquisición de conceptos ad hoc y la formulación de conjeturas para explicar conjuntos de observaciones, incluso quizá un conjunto que contiene un único dato, como mis inducciones sobre los bares italianos o la más sorprendente observación de Sir Alexander Fleming de la destrucción de bacterias en una placa de cultivo –una observación que le llevó al descubrimiento de la penicilina–. todas estas inferencias son falibles, pero podemos ser conscientes de su falibilidad.

* * *

La teoría que voy a esbozar tiene como asunción fundamental que la inducción depende del conocimiento. Para establecer este punto central por sí mismo, podría intentar hacer una inducción sin apoyarse en el conocimiento. Por ejemplo, suponga que sobre la pantalla de su ordenador aparece una extraña mancha de color violeta. No tiene usted ni idea de lo que es. Se acerca de forma ondulante y de repente explota, sin dejar restos. Aparece una nueva mancha violeta, se acerca de forma ondulante y explota. Aparece una tercera mancha, y hace lo mismo. Y ahora aparece una cuarta mancha. ¿Qué espera usted que ocurra? Imagino que usted pensará: debería explotar también. tiene esta idea porque sabe que las manchas anteriores explotaron. Por tanto, aunque empezó sin ningún conocimiento previo sobre las manchas, adquirió enseguida el suficiente como para llevarle a esperar ciertos sucesos. Si cada una de las tres manchas hubiera hecho algo diferente, la idea que usted hubiera tenido sobre lo que podría hacer la siguiente mancha hubiera sido escasa, excepto que quizá haría algo diferente de las anteriores. Del mismo modo, si usted observara que las manchas moradas explotan, las manchas verdes se reducen a nada y las manchas amarillas giran aumentando y aumentando hasta llenar la pantalla, su predicción sobre una mancha dependería de su color. Es usted sensible a la frecuencia con que se producen diferentes sucesos y puede utilizar su conocimiento sobre esas frecuencias para hacer inducciones sobre lo que debería suceder.

La inducción de que una mancha violeta explotará es un ejemplo de lo que los teóricos llaman a veces una inducción por “enumeración”. A medida que encontramos más y más ejemplos del mismo tipo de entidad, como por ejemplo

una mancha morada que explota, es más probable que infiramos que tal vez todas las manchas moradas explotan. Estas inducciones están en un extremo de la escala de homogeneidad. A menudo, las inducciones se refieren a las frecuencias de los sucesos. Por ejemplo, cuando un criptoanalista inglés estaba trabajando en descifrar el código de la máquina Enigma alemana durante la Segunda Guerra Mundial –un problema al que volveré en capítulos posteriores–, sus observaciones le llevaron a inferir que los alemanes estaban transmitiendo al menos cien mensajes al día con una determinada clave para el código.

Muchas inducciones de la vida real dependen del conocimiento previo. Eso reduce nuestra necesidad de realizar múltiples observaciones, y nos capacita para hacer inducciones a partir de sucesos únicos. Si nos enteramos de que Dovis hizo la oferta más baja en el contrato para construir el bloque de apartamentos, entonces podemos inferir que consiguió el contrato. Si nos enteramos de que un hombre fue arrestado a la salida de unos almacenes con ropas por las que no había pagado embutidas en bolsillos secretos de su abrigo, entonces podemos inferir que es un ladrón. Si descubrimos que el motor de arranque de nuestro coche hará funcionar el motor, podemos inferir que la batería está descargada. Sabemos que las ofertas más bajas tienden a conseguir el trabajo, que los ladrones roban ropa de los almacenes escondiéndosela encima, y que los motores de coche a menudo no consiguen arrancar porque la batería está descargada. los experimentos han corroborado nuestra propensión a hacer estas inferencias.

También sabemos que podríamos estar equivocados. Puede que la oferta de Dovis haya sido rechazada porque era demasiado baja para ser factible. Puede que el hombre con las ropas embutidas en sus bolsillos haya sido contratado para evaluar la seguridad del almacén. Puede que el motor de arranque no haya hecho funcionar el motor porque hubiera un fallo en el circuito. Necesitamos más conocimiento para eliminar estas posibilidades: necesitamos evaluar nuestras inducciones. Pero las evaluaciones exigen esfuerzo y pueden no ser necesarias para nuestros objetivos habituales. En la vida cotidiana surgen con tanta frecuencia inducciones específicas porque rara vez tenemos información suficiente como para hacer deducciones útiles. Necesitamos ir más allá de la información que tenemos.

* * *

No existe ningún programa de ordenador comprehensivo para hacer inducciones, aunque los programas simulan diversos tipos de aprendizaje a partir de ejemplos. Para simular las inducciones, es necesario representar el conocimiento y utilizarlo para construir modelos del mundo. Puedo poner un ejemplo de lo que necesita hacer un programa, y de la interacción entre inducción y deducción, con una reconstrucción de mis inferencias iniciales en el bar. Mi primera inferencia fue una inducción “enumerativa”:

Todas las personas que esperan a ser servidas tienen recibos.

El proceso podría basarse en una regla con la forma:

El sujeto a tiene un objeto g.

El sujeto b tiene un objeto g.

El sujeto c tiene un objeto g.

...

Por tanto, todos los sujetos tienen un objeto g.

Después de observar una serie de veces que miembros de un conjunto tienen una

propiedad o relación en común, la regla se pone en marcha y nos permite inferir que todos ellos tienen la propiedad o relación en común. El problema con esta regla es que da lugar a generalizaciones espúreas que ninguno haríamos. Estoy esperando a mi mujer en el aeropuerto. la primera persona que sale del avión es un hombre, y lo mismo la siguiente y la siguiente. Supongamos que las primeras diez personas son todas de sexo masculino. la regla anterior lleva a la generalización de que todas las personas que salgan del avión serán hombres. Por tanto, mi mujer no estaba en el vuelo. la conclusión es absurda, y también lo es la regla sin restricciones.

Lo que se ha pasado por alto es el conocimiento. la base para mi generalización sobre la gente del bar era mi conocimiento de un principio abstracto: los establecimientos del tipo de los restaurantes, bares y cafés tienen un "guion" que su clientela sigue. Sus componentes principales son pedir, consumir, pagar. los dos primeros componentes exigen una cierta rutina de acciones e interacciones con el personal. la rutina puede permitir excepciones o desviaciones. Especifica qué deberíamos hacer por defecto. Pero, cuanto más grande y más ajetreado es el establecimiento, más probable es que exista una rutina estándar, que violamos bajo nuestra propia responsabilidad. Por consiguiente, si al menos algunos de los que esperan a que les sirvan su café tienen recibos, entonces es probable que todos ellos tengan recibos.

Nuestro conocimiento podría formar parte de un sistema que utiliza reglas formales. Sería un equivalente mental a la premisa:

Los establecimientos del tipo de los bares tienen un orden estándar según el cual los clientes piden bebidas, pagan y obtienen recibos.

Se sigue que este bar tiene un orden estándar. Sea cual sea el orden, se aplica a todo el mundo. Mi observación era que algunos clientes que esperaban a pedir en la barra tenían recibos. Y por tanto puedo utilizar las reglas formales para extraer la conclusión: todos los que están esperando a pedir en la barra tienen recibos. Muchas propuestas teóricas sobre la inducción se apoyan en esta idea, y la tabla 13.1 resume algunas reglas formales típicas utilizadas por estos sistemas.

Un inconveniente de las reglas es que no existe forma clara de determinar qué

regla, o qué secuencia de reglas, utilizar con el fin de hacer una inducción.

Cualquier inferencia en la que la conclusión se cumple en, al menos, todas las posibilidades consistentes con las premisas es una deducción válida; a la inversa, cualquier inferencia en la que la conclusión se cumple en no todas las posibilidades consistentes con las premisas es una inducción. la primera regla de la tabla 13.1 elimina posibles sujetos –aquellos que no tienen la propiedad importante–. El efecto de cada una de las reglas restantes se alcanza con una única operación sobre los modelos, a saber, la eliminación de uno de los modelos para una posibilidad. Por ejemplo, la proposición:

Si has pagado y tienes un recibo entonces te servirán

es compatible con un conjunto de modelos que incluyen la posibilidad:

pagado no recibo no servido

Tabla 13.1. Reglas formales de inferencia para generalizaciones inductivas

Muchos teóricos han abogado por las reglas formales para la generalización inductiva. Ese tipo de reglas incluye las siguientes:

1. El ítem 1 tiene la propiedad P; lo mismo sucede con los ítems 2, 3, 4, etc.

Por tanto, todos los ítems tienen la propiedad P.

2. Si A y B entonces C.

Por tanto, si A entonces C.

3. Si A entonces (B o C).

Por tanto, si A entonces C.

4. Si A y B entonces C; y si A y D entonces C.

Por tanto, si A entonces C.

5. Si A entonces C.

Por tanto, si A o B entonces C.

6. Si A entonces C.

Por tanto, si A entonces (B y C).

Un ejemplo de la utilización de la regla 2 sería argumentar: Si has pagado y tienes un recibo entonces te servirán; por tanto, si has pegado entonces te servirán. El efecto de todas estas reglas es eliminar posibilidades.

Si este modelo se elimina, el restante conjunto de modelos es equivalente a la conclusión:

Si has pagado entonces te servirán.

Por tanto, la inducción es equivalente a la regla 2 de la tabla. la operación de añadir información a un modelo también puede enriquecerlo al introducir una nueva relación.

Otro inconveniente de las reglas formales es que las premisas pueden utilizar términos diferentes a los de la representación mental del conocimiento, y por tanto las premisas no pueden desencadenar el conocimiento. Una solución es introducir en el conocimiento principios adicionales que permitan a las premisas enlazar con el conocimiento. Estos postulados del significado incluirían los axiomas:

Pagas significa das dinero a alguien por algo.

Un pedido significa una petición de mercancías o servicios.

Un cliente significa una persona que paga por mercancías o servicios.

Tal vez podría hacerse funcionar el sistema, pero los postulados del significado tienen varias dificultades serias, que describí en capítulos anteriores. Una alternativa es utilizar modelos para representar el conocimiento. tienen la ventaja de que se puede encajar en ellos la percepción visual –la base de mis observaciones en el bar– sin la intermediación del lenguaje. En la deducción, nos concentramos en lo que aparece de forma explícita en nuestros modelos, y por tanto podemos dejar de hacer ciertas deducciones. En la inducción, también nos centramos en lo que aparece de forma explícita en nuestros modelos, y por tanto tratamos de encontrar pruebas que corroboren nuestras hipótesis inductivas. No buscamos ejemplos negativos de nuestras hipótesis, y los encontramos solo por causalidad.

La etapa más misteriosa de todo mi proceso de pensamiento fue el reconocimiento de que el trocito de papel que sostenía el hombre que estaba a mi lado era un recibo. Es también una inducción. Pero, a pesar de haber investigado

el proceso durante más de un siglo, no se sabe bien cómo reconocemos los objetos.

* * *

David Hume hizo que los filósofos se preocuparan por la justificación racional de la inducción. En realidad, como él sostuvo, no tiene ninguna (que a su vez no dependa de la inducción). Pero la deducción tampoco se puede justificar sin deducción. la inducción es solamente algo que los animales, incluidos los seres humanos, hacen para hacer posible la vida. Más recientemente, los filósofos se han preocupado por qué propiedades de los objetos garantizan las inferencias inductivas. la respuesta se apoya en el conocimiento: no inferimos que todos los pasajeros de un avión son hombres porque lo sean los diez primeros que salen del avión. Sabemos que esta observación no descarta la posibilidad de que haya un pasajero que sea mujer. Por el contrario, una experiencia me bastó para inferir el orden para pagar y pedir en los bares italianos que tenían un cajero.

El número de posibles inducciones que podemos hacer sobre un conjunto finito de datos es demasiado amplio para ser abarcable. El conocimiento nos restringe, sin embargo. Su abarcabilidad relativa es en sí misma una restricción. la información sobre un tipo de suceso es más accesible que la información sobre otro tipo de suceso, y por tanto inferimos que el primer tipo de suceso es más probable que el segundo. Puede haber también otras restricciones generales. Una es formular la hipótesis más informativa que sea consistente con las pruebas, es decir, la hipótesis que admite el menor número de posibilidades. Esta restricción es crucial cuando solo podemos observar ejemplos positivos de un concepto y no tenemos un profesor que señale ejemplos negativos para ayudarnos. Por ejemplo, si encontramos pacientes que sufren de diarrea, vómitos, baja presión sanguínea y espasmos musculares, entonces la hipótesis más específica es que están sufriendo una única enfermedad con esos síntomas. Si ahora encontramos varios pacientes similares que tienen todos esos síntomas excepto los espasmos musculares, probablemente nuestra hipótesis inicial era demasiado específica. Pero si hubiéramos empezado con la hipótesis de que el primer paciente estaba sufriendo varias enfermedades simultáneas, no hubiéramos tenido una razón para cambiar nuestra idea a la vista de los posteriores pacientes.

Otra restricción general es la similitud. Si una mancha es similar a una que explotó, entonces también puede explotar. Como señaló Hume, causas similares tienen efectos similares, y este principio parece estar detrás de muchas de nuestras inducciones. Es decir, inferimos que un suceso es similar a otro, y esperamos que tenga consecuencias similares. lo que es problemática es la base que hace que juzguemos que existe similitud. Algunos casos son lo bastante simples: los dos sucesos son difíciles de distinguir. Pero no existe ningún programa de ordenador general que evalúe la semejanza entre las cosas. Parte del problema es que entre dos entidades cualquiera elegidas al azar casi siempre podemos encontrar alguna semejanza. ¿En qué son similares un piano y una banana? Bien, ambos son objetos físicos; ambos tienen formas características; ambos tienen un interior y un exterior; ambos gustan a la gente; ambos deben su existencia a los árboles; etc. la moraleja es que usamos la similitud como base para las inducciones, pero no es una noción transparente, y con frecuencia depende de nuestro conocimiento de las entidades. Un murciélago parece un pájaro, pero un murciélago es un mamífero y no es un pájaro. Un pastor alemán no se parece a un chihuahua, pero ambos son perros.

* * *

He aquí un hecho que puede que usted no conozca:

Las mujeres de inteligencia excepcional tienden a casarse con hombres menos inteligentes.

¿Por qué? Si usted comenta esto en una conversación, la gente creará todo tipo de explicaciones improvisadas al respecto –tal vez a las mujeres inteligentes no les gustan los hombres inteligentes, tal vez se sienten amenazadas por ellos, tal vez les gusta dominar a sus maridos, etc.– Hay un puñado de problemas similares. los padres brillantes tienden a tener hijos menos brillantes. los atletas que tienen una temporada de grandes éxitos tienden a tener a continuación una temporada menos estelar. Una comida terrible en un restaurante tiende a ser

seguida por una comida menos terrible la siguiente vez que uno come allí. De nuevo, mencione uno de estos hechos en la conversación, y su amigos le asombrarán con sus ingeniosas explicaciones. Sin embargo, no se deje ganar por sus ideas. Una realidad estadística, “la regresión a la media”, puede explicar todos estos fenómenos. las observaciones en el extremo de una distribución son raras. Solo hay unos cuantos sujetos de inteligencia excepcional. Muchos más están próximos al término medio, es decir, a la media. Por consiguiente, las posibilidades son que un sujeto excepcional se encuentre y se case con alguien más próximo a la media o al término medio. los atletas que tienen un año excepcional es probable que lo logren como resultado tanto del talento como de la suerte. Al año siguiente su actuación tiende a volver a su término medio a lo largo de los años. El concepto de regresión a la media es difícil. Es un conocimiento técnico perteneciente a la estadística. Podemos entenderlo cuando nos lo explican. Pero sus ejemplos no nos resultan tan fáciles de identificar. Por el contrario, nuestra propensión a crear explicaciones nos lleva a violar las normas de la probabilidad.

* * *

¿Cómo hacemos inducciones? la teoría esbozada en este capítulo no considera que la inducción y la deducción sean tan diferentes como plantea la lógica. Cuando hacemos deducciones, a menudo traspasamos la frontera con la inducción porque nos basamos en conocimiento que va más allá de las premisas. El conocimiento es falible: el precio de la inducción es la no validez. Normalmente nos centramos en los éxitos de la inducción y las imperfecciones menores que aportan claves acerca de su naturaleza. Y pasamos por alto sus catástrofes –las modas pasajeras de la pseudociencia, las supersticiones de la vida diaria y las desastrosas inducciones que llevan a sucesos como el hundimiento del The Herald of Free Enterprise–. El origen de estos errores está en la inducción humana. Como ya mencioné anteriormente, su heurística, o reglas generales, son lo que los que estudian otras culturas llaman pensamiento “mágico”.

CAPÍTULO 14

El método de Sherlock Holmes: la abducción

Un médico escocés entrevistando a un paciente hace una serie de inferencias acertadas:

Bien, amigo mío, usted ha servido en el ejército.

Sí, señor.

No hace mucho que fue dado de baja.

No, señor.

¿Un regimiento de las tierras altas?

Sí, señor.

¿Suboficial?

Sí, señor.

¿Destinado en Barbados?

Sí, señor.

Estas inferencias pueden recordarle a usted al gran detective de ficción de Conan Doyle, Sherlock Holmes. En la historia La liga de los pelirrojos, Holmes deja atónito a su amigo el Dr. Watson. Ha venido a consultarles un visitante, y Holmes señala: “Aparte de los hechos evidentes de que en alguna época ha realizado trabajos manuales, que toma rapé, que es masón, que ha estado en China y que últimamente ha escrito muchísimo, soy incapaz de deducir nada más.” la similitud entre Holmes y el doctor escocés no es un accidente. Cuando Conan Doyle era estudiante de medicina en Edimburgo, había trabajado como ayudante del doctor que mantuvo ese diálogo, Joseph Bell. Y más tarde describió cómo hizo Bell su asombrosa cadena de inferencias:

“Vean ustedes, señores”, explicaría, “el hombre era un hombre respetuoso, pero no se quitó el sombrero. En el ejército no lo hacen, pero si llevase mucho tiempo de baja habría aprendido las costumbres civiles. tiene un aire de autoridad y obviamente es escocés. En cuanto a las Barbados, su dolencia es elefantiasis, que se da en las Indias occidentales y no en Gran Bretaña”.

Para su audiencia de Watsons todo parecía bastante milagroso hasta que lo explicó y entonces se convirtió en bastante sencillo. No es asombroso que después del estudio de un personaje como este utilizara y ampliara sus métodos cuando en mi vida posterior intenté construir un detective científico que resolviera casos por sus propios méritos y no por la estupidez del criminal.

Así que, cuando Holmes explicó la base de sus inferencias, el visitante señala, “Al principio me pareció que había hecho usted algo muy inteligente, pero ahora me doy cuenta de que, después de todo, no tiene ningún mérito”.

Holmes decía que su método era deductivo, y que sus deducciones eran tan infalibles como las pruebas de Euclides. Estaba equivocado, infirió, por ejemplo, que el visitante era francmasón porque llevaba una insignia con un arco y un compás. ¿Se deduce necesariamente que un hombre que lleva este alfiler es un francmasón? Por supuesto que no. Podía llevarla con la intención de engañar, o sin darse cuenta siquiera de su significación. Algunas veces echo de menos un anti-Holmes que anuncie sus “brillantes” deducciones con la misma indiferencia que el gran detective, pero que se equivoque en todos los casos.

Sin embargo, Sherlock Holmes y su modelo, Joseph Bell, están razonando. Por tanto, ¿cuál es su método? En otra historia, La aventura de la caja de cartón, Holmes da una versión más exacta de su pensamiento. “El caso es uno en el que... nos habíamos visto obligados a razonar retrocediendo desde los efectos a las causas.” Este tipo de razonamiento da lugar a explicaciones. ¿Por qué lleva el hombre una insignia con un arco y un compás? Respuesta: porque es un francmasón. ¿Por qué no se quita el sombrero? Respuesta: porque ha sido dado de baja del ejército recientemente. ¿Por qué ha contraído elefantiasis? Respuesta: porque estaba destinado en Barbados, donde era endémica. Por supuesto, Holmes y Bell tenían que percibir esas pistas antes de poder explicarlas. Pero su método es abductivo y falible.

En el capítulo anterior, describí la interrelación entre inducción y abducción. Resolví el problema de cómo conseguir un capuchino en un bar italiano: uno paga primero y luego da el recibo al camarero. Esta inducción parecía aplicarse a los bares italianos en general porque fui capaz de hacer una abducción que la explicara: los camareros estaban demasiado ocupados haciendo cafés como para ocuparse de la caja registradora. Una explicación refuerza nuestra creencia en una inducción. Peirce acuñó el término “abducción” para referirse al proceso inferencial que da lugar a explicaciones causales, porque, como Sherlock Holmes, razonamos retrocediendo desde un efecto a la causa que lo explica. Sin embargo, yo me he apropiado del término para referirme a inferencias que dan lugar a cualquier tipo de explicación, y mi propósito es intentar describir el proceso mental que subyace en ellas.

* * *

Los filósofos han sostenido que la abducción debe dar lugar a la mejor explicación posible de los hechos. Esto parece loable. El problema es describir qué se considera como mejor. Aparecen en juego varios factores. Una buena explicación debe decir la verdad. Una que defienda que los bares tienen cajeros porque los camareros están demasiado ocupados para enfrentarse a los pagos es falsa si, pongamos por caso, los camareros tienen mucho tiempo pero no quieren manejar dinero.

Un segundo factor en una buena explicación es la parsimonia. Desgraciadamente, no es fácil de evaluar. Si contamos el número de palabras de una descripción, entonces depende de las palabras concretas de nuestro vocabulario. Variará también de un idioma a otro. Podríamos pensar en definir la parsimonia en función del programa de ordenador más corto que da lugar a la explicación (en algún lenguaje universal de programación). Pero esta idea conduce a problemas aún peores, que he dejado para las notas del capítulo. Quizá el mejor camino es pensar en la parsimonia como hizo el filósofo medieval William de ockham. Se supone que declaró: “Entia non sunt multiplicanda praeter necessitatem.” Es decir, “las entidades no tienen que multiplicarse sin necesidad”. El principio se conoce como la navaja de ockham, y con frecuencia guía las explicaciones científicas. Examinamos el número de diferentes tipos de entidad en una explicación y tal vez el número y complejidad de las relaciones entre ellas.

Un tercer factor en una buena explicación es el número de observaciones que cubre. Mi abducción no explicaba demasiado, más allá del orden de los asuntos en los bares italianos. Por el contrario, la explicación científica de unas observaciones a menudo supera a estas para explicar otras observaciones, e incluso predecirlas. Ptolomeo explicó los movimientos de los planetas pero Newton además de eso explicó sucesos locales en la tierra, como la caída de una manzana. la explicación de Newton es mejor que la de Ptolomeo, porque está más cerca de la verdad, es más parsimoniosa y es de mayor ámbito. El resultado final es que una buena explicación es deseable, imposible de probar y es una cuestión de juicio informal.

* * *

¿Entiende usted cómo funcionan las máquinas de coser? Por supuesto que sí. Pero, veamos lo bien que las entiende. Usted sabe que la máquina tiene una aguja y un hilo y que los empuja a través de los trozos de tela para que queden cosidos uno a otro. Pero, ¿qué ocurre luego? A través de la tela sale un lazo de hilo y, de alguna manera, otro hilo pasa a través del lazo, y se tensa contra la parte inferior de la tela. Pero, ¿cómo puede pasar un hilo a través del lazo? tal vez usted (y yo) nos hemos dado cuenta ahora de que sobreestimábamos nuestra comprensión de las máquinas de coser. Es un estado mental que se produce con tanta frecuencia que los psicólogos se refieren a él como la ilusión de la profundidad de la comprensión. No cometemos el error equivalente respecto a nuestro conocimiento de los hechos. Sin embargo, a menudo pensamos que entendemos algo solo para descubrir, al tratar de explicarlo, que tenemos como máximo una idea superficial. Para comprender profundamente algo, descríbalo, enséñelo y luego investíguelo.

Como han mostrado los estudios, cuando explicamos cómo funciona un aparato, primero describimos cómo funciona a un cierto nivel, y luego esbozamos el mecanismo subyacente. Podemos progresar hasta un nivel más profundo, y describir el funcionamiento de cada componente del mecanismo, y luego sus mecanismos, etc. En el caso de una máquina de coser, hay una bobina de hilo que gira debajo del plato en el que descansa la prenda, y cuando el ojo de la aguja empuja el lazo de hilo a través de la prenda, el hilo de la bobina se hace pasar a través del lazo. Pero, ¿cómo? Este es precisamente su truco. la bobina lleva enrollado un hilo largo, un extremo de este hilo ya se ha cosido en la prenda, pero el otro extremo del hilo está enrollado en la bobina. Es este extremo libre el que queremos meter a través del lazo del hilo superior. la máquina no puede empujar la bobina a través del lazo. Por tanto, en lugar de eso, tira del lazo en torno a la bobina, que está girando verticalmente. El lazo saca luego hacia arriba el hilo de la bobina tensándolo contra el lado inferior de la prenda. Y ¿cómo hace todo eso la máquina? la respuesta es que cuando el lazo pasa a través de la prenda, se engancha en una muesca de la bobina. A medida que la bobina gira, la muesca tira del lazo de hilo en torno a ella. Para entonces, la aguja se está moviendo hacia arriba y así el lazo asciende con el hilo de la bobina en su interior tensado contra la prenda. Ya puede repetirse el ciclo completo. Como puede usted ver, mi descripción ha seguido la receta general de

explicar asuntos a un nivel cada vez más profundo. En caso de que tenga dificultad para seguirla, la Figura 14.1 le ofrece seis “instantáneas” de la secuencia para ayudarle a entender cómo funciona el mecanismo.

* * *

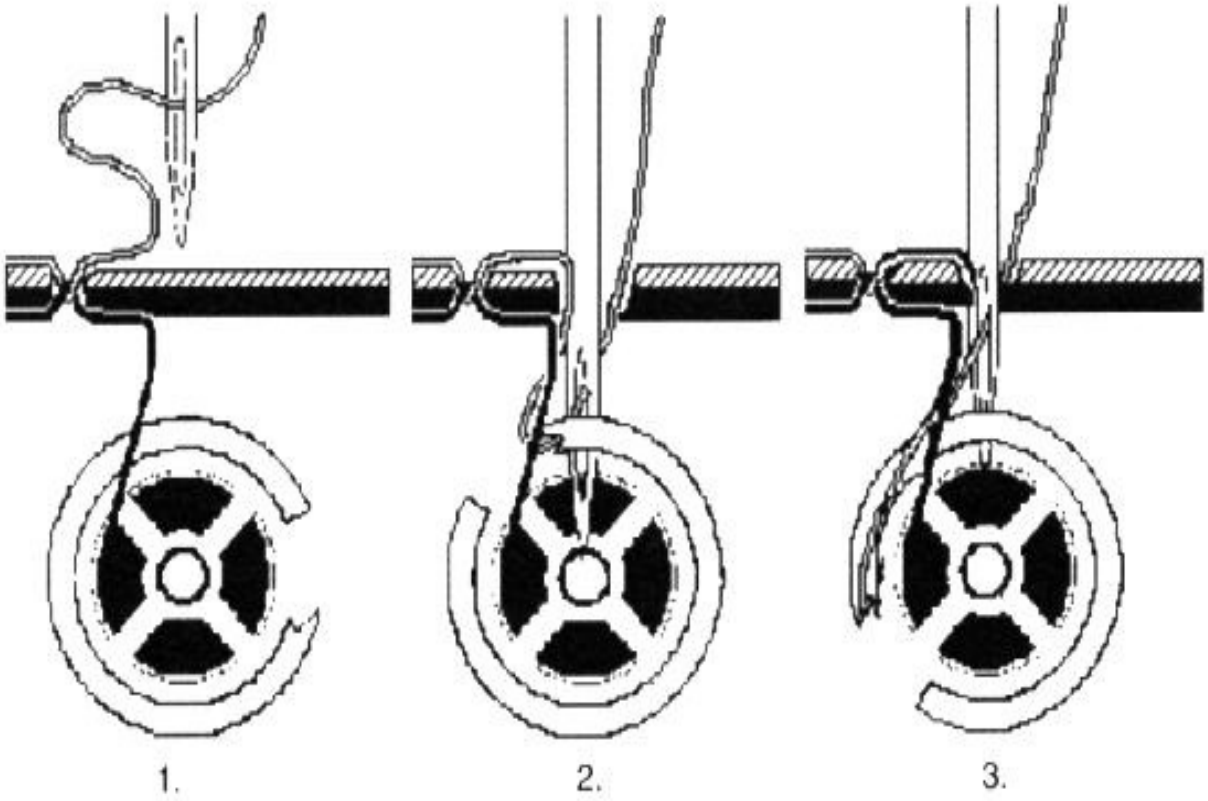
La anterior descripción de una máquina de coser es causal. Explica cómo causa la máquina un paso del proceso, el cual da lugar a otro, etc. Si no estuviéramos familiarizados con lo que hace una máquina de coser, nos sería difícil crear esta secuencia –habríamos tenido que reinventar la máquina de coser–. Muchas explicaciones sencillas en la vida, sin embargo, se pueden construir a partir de nuestro conocimiento general sobre las causas y sus efectos. la capacidad para imaginar posibles causas está en el centro de las explicaciones. Utilizamos nuestro conocimiento general para recuperar o construir cadenas causales. todas las culturas parecen utilizar explicaciones causales, y muchas culturas, incluida la nuestra, les dan un gran valor. los occidentales, sin embargo, tendemos a explicar las acciones de la gente desde el punto de vista de sus intenciones y capacidades, mientras que los orientales tienden a explicarlas desde el punto de vista del contexto social.

Una explicación verosímil de un hecho físico debería ser de tipo causal, y, dada la opción entre una deducción y una explicación causal, tanto los occidentales como los orientales tienen tendencia hacia esta última. He aquí un ejemplo sencillo de esta diferencia. Supongamos que yo le digo: si un piloto cae sin paracaídas de un avión entonces el piloto muere. Ahora bien, este piloto no murió. Y le pregunto: ¿Por qué no? la mayoría de los participantes en un experimento respondieron con estos tipos de explicación:

El avión estaba en el suelo y por eso él (sic) no cayó desde muy arriba.

Cayó en un gruesa capa de nieve acumulada por el viento y por eso su caída se interrumpió.

El piloto ya estaba muerto y por eso la caída no le mató.


1.
2.
3.

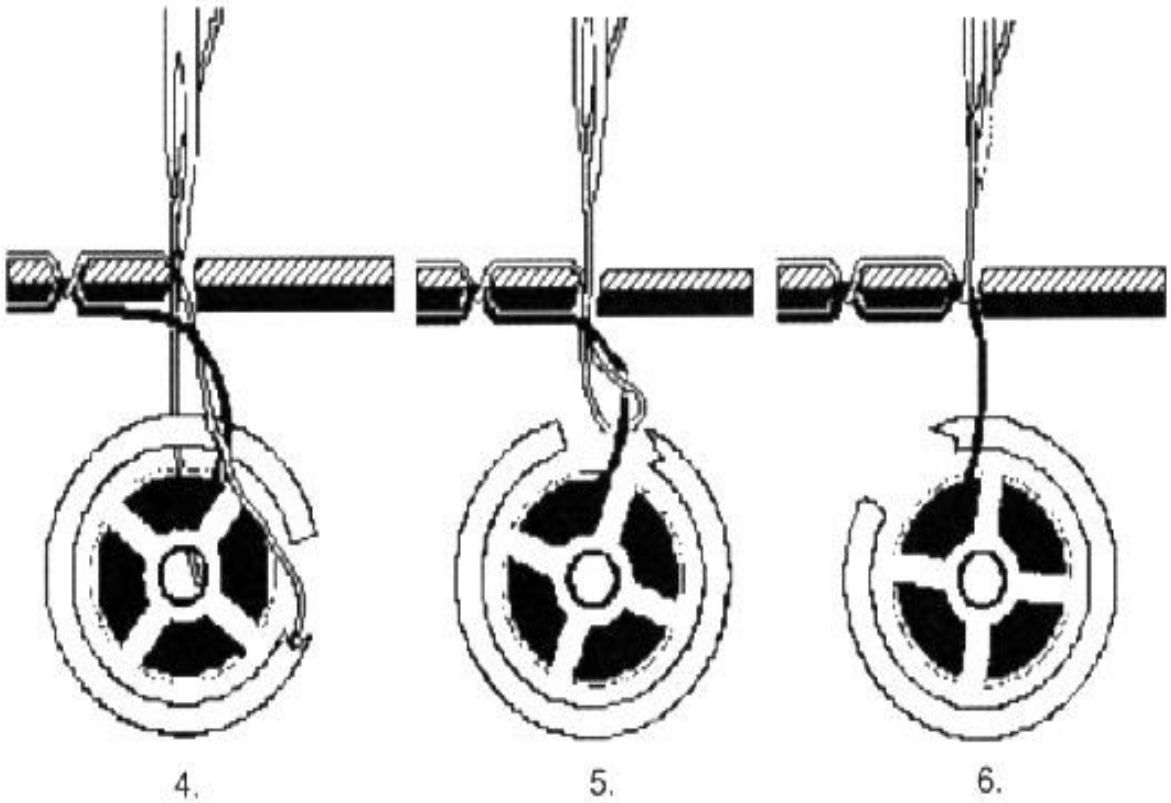

4.
5.
6.

Fig. 14.1. Seis instantáneas que ilustran cómo funciona una máquina de coser.

Solo unos cuantos respondieron con una deducción:

El piloto no cayó sin paracaídas de un avión.

Puede que usted haya inferido que realmente cayó sin paracaídas. la forma en que yo formulé la pregunta puede haberle llevado a usted a hacer esa inferencia. No obstante, si yo le dijera que realmente sí cayó sin paracaídas, usted no tendría dificultad para plantear explicaciones causales de por qué no murió, una vez que le he impedido ofrecer la versión deductiva.

La mejor clase de explicación para los hechos es causal. Pero, ¿cuál es la longitud óptima de la cadena? Supongamos que usted se entera de que un conocido mío ha sufrido un accidente de coche, y me pide una explicación. Podría decirle, perdió el control del coche. Pero esta explicación parece ad hoc y sin motivo: sigue necesitando una explicación. Una explicación más verosímil proporciona un motivo en forma de causa: iba conduciendo demasiado rápido por una carretera helada y perdió el control del coche. Por supuesto, usted también podría pedir una explicación para esta causa, pero a medida que una cadena causal crece, resulta más improbable. Ciertamente, la explicación precedente es ya más improbable que la explicación mínima de que el conductor perdió el control del coche. No obstante, lo óptimo parece ser una causa y un efecto. la causa proporciona motivación para el efecto; el efecto explica el suceso en cuestión; y la explicación no es tan larga como para parecer improbable.

¿Cómo creamos explicaciones causales? Algunas están en la memoria a largo plazo, esperando que las utilicemos. otras no lo están, y tenemos que crearlas. Conocemos muchas relaciones causales elementales, que podrían representarse en el conocimiento como modelos explícitos de conjuntos de posibilidades, y

podríamos hilvanar esos eslabones para formar una nueva cadena de sucesos.

Supongamos que tenemos que explicar por qué un hombre ha saltado a una piscina completamente vestido. Puede que no conozcamos ninguna razón para esta acción. Pero sabemos que es estúpida y podemos recuperar una causa para acciones estúpidas, como por ejemplo la ebriedad. Y por tanto podemos inferir: quizá el hombre estaba borracho. Esta posibilidad a su vez puede desencadenar una nueva causa incluida en nuestro conocimiento general: quizá estaba borracho porque su novia le había dejado plantado. Podríamos seguir alargando la cadena causal de esa manera. Estamos construyendo una historia:

Quizá su novia le había plantado porque pensaba que le había sido infiel, aunque no lo había sido, y por eso desesperado bebió demasiado y como resultado saltó a una piscina totalmente vestido.

Estas cadenas pueden ser originales en el sentido de que nunca antes hemos pensado en ellas. Recuperamos los eslabones de la memoria, pero en cada punto pueden venirnos a la mente varias posibilidades, y seleccionamos aquellas que parecen apropiadas para construir la cadena. Nuestro razonamiento inconsciente está haciendo una secuencia de transiciones de Markov del tipo de las que describí en el Capítulo 5.

Escribí un programa de ordenador que construye cadenas de relaciones causales por ese sistema. Se basa en la asunción de que nuestro conocimiento de las posibilidades se representa en modelos de las causas y sus efectos –algunas pruebas psicológicas corroboran este orden–. la entrada [input] del programa es un modelo de un suceso. El programa utiliza el contenido de este modelo para buscar en su base de conocimiento y, si es posible, desencadenar una posibilidad relevante que lo explique. Si la búsqueda tiene éxito, el programa utiliza a su vez esa posibilidad para buscar de nuevo en su base de conocimiento, y, si es posible, desencadenar una posibilidad que causa el nuevo efecto. Por esta vía, el programa construye una cadena causal que funciona hacia atrás partiendo del suceso hasta su explicación causal. Su plan es análogo a la forma en que Conan Doyle escribió muchas de sus historias. Pensaba en un crimen y luego, a partir de él, construía una cadena que retrocedía hasta un suceso concreto y

desconcertante que exigía que Sherlock Holmes lo desentrañara. ¿Por qué, por ejemplo, iba alguien a poner anuncios para que hombres pelirrojos solicitaran unirse a una liga que les diera derecho a recibir una gran cantidad por servicios puramente nominales? El lector se engancha lo mismo que se enganchó Holmes.

La gente es imaginativa al construir explicaciones. Un experimento mostró que su capacidad superaba a cualquier programa de ordenador existente. Podían explicar conjunciones arbitrarias de sucesos. Se les daban parejas de oraciones tomadas al azar de diferentes historias, p. ej.:

John llegó hasta una tienda que vendía televisores.

Celia acababa de perforarse las orejas.

En otra situación del experimento, se modificaron las oraciones para asegurar que se hacía referencia al mismo sujeto en ambas oraciones del par:

Celia llegó hasta una tienda que vendía televisores.

Ella acababa de perforarse las orejas.

Los participantes encontraban un determinado contenido una sola vez, y tenían que explicar qué estaba sucediendo en cada situación. Iban más allá de la información que se les daba para dar una versión de lo que estaba pasando. Proponían, por ejemplo, que Celia estaba recibiendo señales en sus pendientes y buscaba la tienda de televisores para investigar, que llevaba pendientes nuevos y quería verse en un circuito cerrado de televisión, que había ganado una apuesta por haberse perforado las orejas e iba a gastar el dinero en un televisor nuevo,

etc. los participantes fueron casi igual de ingeniosos con las oraciones que no se referían a un único sujeto. Pero a veces no supieron qué explicación dar. Este fallo es al que me refiero como el efecto “Mary Celeste”. El ejemplo epónimo se basa en un caso histórico: creemos que si abordamos un barco en el mar, la tripulación estará allí. Abordamos el Mary Celeste. Descubrimos que la tripulación no está a bordo. Y, como los marineros que lo abordaron originalmente, somos incapaces de inferir qué les ha sucedido. Este efecto Mary Celeste muestra que nuestro razonamiento cotidiano no siempre tiene éxito.

* * *

Algunos psicólogos sostienen que nuestro razonamiento cotidiano depende de un procedimiento abarcable, dado lo rápidamente que ideamos hipótesis explicativas. El argumento parece verosímil, pero es imposible evaluar la eficiencia de un procedimiento a menos que uno sepa lo que se supone que computa. Inferir que Celia quería ver sus pendientes en televisión parece una explicación satisfactoria al problema anterior, pero consideremos esta respuesta:

Celia quería comprar un televisor.

¿Es satisfactoria? No dice nada sobre la perforación de las orejas, y quizá por eso no es una explicación adecuada. Y qué hay de la respuesta Mary Celeste:

No puedo explicar lo que está pasando.

¿Cuenta como satisfactoria? Si lo hace, entonces el proceso es eficiente en un sentido trivial: sabemos cuándo no podemos explicar un cierto suceso. Si no lo hace, entonces quizá nuestro razonamiento cotidiano, después de todo, no es tan

eficiente porque no siempre da lugar a una explicación.

* * *

La abducción confía en otra fuente de explicaciones: la analogía. Como ejemplo consideremos este problema:

El enemigo de tom, Jerry, fue apuñalado hasta morir en un cine durante la proyección vespertina de Bambi. Cuando tuvo lugar el asesinato tom estaba en un tren a Filadelfia. ¿Qué se sigue?

Una inducción verosímil es que tom es inocente. Representamos las premisas en un modelo mental con el cine en un lugar y el tren en otro. Jerry está en el cine; tom está en el tren. Por tanto, tom no podría haber apuñalado a Jerry. En un experimento con una serie de problemas de este tipo, todos los participantes inferían que tom era inocente. Entonces el experimentador les decía, “No, no es eso”. luego les preguntaba: “¿Puede pensar usted en alguna otra posibilidad?” Así que, ¿puede pensar usted en alguna otra posibilidad?

Su tarea es hacer una abducción que explique cómo, a pesar de lo dicho, tom podría ser culpable del asesinato de Jerry. Una maniobra inicial importante es una modificación espacial del modelo con el fin de satisfacer las restricciones para que tom apuñale a Jerry. Puede usted recolocar el cine de modo que esté en el tren. No hay nada en las premisas que descarte esa posibilidad. Por tanto, usted sugiere: el cine estaba en el tren. No, dice el experimentador, no es eso. Una maniobra espacial similar permite al tren pasar por el cine –es un cine al aire libre–. tom se asoma por la ventana, y está armado con un cuchillo muy largo. No, dice una vez más el experimentador, no es eso. (A estas alturas, usted puede haber inducido que va a dar esta respuesta a todas sus ideas.) En el experimento variaba la capacidad de los participantes para pensar en posibles situaciones. Pero las que creaban los distintos participantes tendían a aparecer aproximadamente en el mismo orden. Por ejemplo, después de que habían

inferido que tom era inocente, tendían a hacer manipulaciones espaciales del tipo de las que acabo de describir.

¿Qué hacer a continuación? Su modelo inicial de tom y Jerry era correcto: estaban en sitios diferentes. Por tanto, ¿cómo podría tom, a pesar de todo, ser culpable del asesinato? En este punto, se necesita un insight. Aquí, puede ayudar una analogía. Puede pensar en una que da lugar al insight. Pero, en caso de que esté usted atascado, déjeme sugerir una. tengo un pequeño jardín unido a mi casa, pero no tengo tiempo suficiente para cuidarlo y por ello empleo a un jardinero que lo atiende por mí. ¿Sugiere esto algunas posibilidades?

Con suerte, usted debería haber tenido la idea de que quizá tom pagó a un cómplice por cometer el asesinato (exactamente igual que yo pago a mi jardinero por segar mi césped). No, dice el experimentador, no es eso, ¿puede usted pensar en algunas otras posibilidades? Mucha gente está dispuesta a abandonar en torno a este punto. Pero, debería usted seguir, porque hay otras posibilidades.

Un cómplice permite a un criminal actuar a distancia, y usted necesita persistir con esta idea. los participantes que perseveran idean nuevas posibilidades para una acción a distancia:

Tom fijó un cuchillo con muelle en el asiento del cine.

Colgó un cuchillo dentro de un bloque de hielo sobre el asiento de Jerry.

Utilizó un robot con radio-control. (los ingenieros tienden a pensar en esta idea.)

Utilizó psicocinesis. (No, dice el experimentador, no es eso.)

Durante los años que el difunto John Frankenheimer, el director de cine, rechazó

que su obra maestra, The Manchurian Candidate, se exhibiera en los cines, solo una persona de los varios cientos a los que les pasé este problema sugirió que tom podría haber producido a Jerry una sugestión posthipnótica para que se apuñalara a sí mismo. (Sucedió en Suecia, y la persona que pensó en esta idea, como descubrí más tarde, era una princesa sueca.) The Manchurian Candidate muestra al psicólogo más siniestro de toda la cinematografía, un hombre del "instituto Pavlov" que utiliza la sugestión posthipnótica como una técnica de asesinato. Después de que Frankenheimer permitiera que la película volviera a exhibirse en los cines, esta respuesta al problema se hizo más frecuente, al menos una persona del público pensaba en ella cada vez que yo lo planteaba. Quizá la película era más accesible como fuente de una analogía.

¿En qué momento se puede inferir que no existen otras posibilidades y que tom es inocente? ¡Ay de tom!, nunca se podrá llegar a esa conclusión. No es deducible: uno nunca puede estar seguro de que ha eliminado todas las posibilidades en las que tom es culpable. Ciertamente, los problemas abductivos parecen más próximos a los de la vida real que los que aparecen en los libros de texto de lógica. los únicos límites a las posibles situaciones son los límites de su imaginación. He oído la mayoría de las posibilidades. Pero, ¿quién sabe? la próxima persona a la que plantee el problema puede crear una solución original basada en una nueva analogía. Ahora que lo pienso, ¡nadie ha sugerido el viaje en el tiempo!

* * *

Existen varias teorías psicológicas sobre la analogía, y en términos generales están de acuerdo en que una vez que hemos encontrado una analogía, transferimos algunos de sus componentes al problema que estamos tratando de resolver. Su problema era mostrar que tom es culpable del asesinato de Jerry, y mi analogía era que empleo a un jardinero para cortar el césped. Por tanto, podemos empezar con estas correspondencias entre la analogía y el problema:

Phil ————— tom

|

emplea

|

jardinero

|

para cortar

|

el césped ——— Jerry

Deberíamos ser capaces de completar las conexiones para plantear una idea nueva que haga culpable a tom:

Phil ————— tom

| |

emplea emplea

| |

jardinero cómplice

| |

para cortar para asesinar a

| |

el césped ——— Jerry

Una vez que hemos pensado en una analogía fructífera, no deberíamos tener muchas dificultades para hacer las conexiones correctas. El problema difícil tanto en la teoría como en la práctica es encontrar la analogía.

Parece que nos diferenciamos entre nosotros en la capacidad para plantear analogías apropiadas. Hay quien tiene una habilidad especial para ello. Wilbur Wright, del que hablaré más adelante, extrajo varias analogías profundas entre la

bicicleta y una máquina voladora más pesada que el aire. Ciertamente, es extraordinario lo válidas que fueron sus analogías, al menos para la primitiva aviación. otra fuente para las analogías de Wilbur fue el vuelo de los pájaros. Sostenía que para que un avión girara, necesitaba hacer lo que hacen algunos pájaros, a saber, retorcer sus alas en direcciones opuestas. Pero, ¿cómo podría hacerse la torsión sin comprometer la estabilidad estructural de las alas? Cuando estaba hablando con un cliente en su tienda de bicicletas, tuvo un insight que es esclarecedor. Sostenía en las manos una caja estrecha y larga de una cámara de bicicleta y por casualidad retorció un extremo de la caja en una dirección y el otro extremo en la dirección opuesta. Inmediatamente vio la analogía con un biplano. Dependía de la visualización de las partes superior e inferior de la caja retorcida como las alas de un biplano. Pero también sintió las consecuencias: la caja retorcida mantenía su rigidez lateral.

Los programas de ordenador que extraen analogías representan el conocimiento en términos lingüísticos. Para representar la idea de que retorcer un ala la hace girar, utilizan representaciones de este tipo, que yo he simplificado un poco:

(causa (retuerce el ala)(gira el ala))

El único significado que tienen estas palabras para los programas es que encajan las mismas palabras en otros datos. Así, la utilización del inglés, frente al latín o los numerales, es por comodidad de quien programa el ordenador. las representaciones verbales serían verosímiles si pensar fuera solamente hablar con uno mismo –una opinión que fue popular en un tiempo entre los conductistas–. El difunto Richard Feynman fue sacado de su error siendo estudiante cuando un amigo escéptico le preguntó qué palabra utilizaría para describir la forma de un cigüeñal. la imaginación visual no es manipulación verbal. Pero tampoco es una manipulación de imágenes bidimensionales. Como sostuve en la Parte I del libro, las representaciones mentales del mundo a menudo son modelos, y la analogía de Wilbur Wright dependía de la manipulación de un modelo tridimensional. El sistema visual construye un modelo mental para la caja. Este modelo le sirve de analogía, y las superficies superior e inferior de la caja se enlazan con un modelo mental de un biplano. “Vemos” las superficies superior e inferior de la caja como alas. Retorcer la caja

es similar a retorcer las alas. Y el sentido cinestésico de torsión y su falta de efecto en la estabilidad lateral de la caja se puede enlazar con el modelo de las alas. Ningún programa de ordenador actual que utilice analogías puede llevar a cabo este proceso de formar analogías entre modelos dinámicos tridimensionales, pero esa era una de las habilidades fundamentales de los hermanos Wright. A menudo, las analogías se producen entre modelos del mundo más que entre proposiciones. El mismo Feynman informa de un método de aplicación general para la construcción de analogías a partir de modelos:

Tenía un esquema, que sigo usando hasta hoy cuando alguien está explicando algo que estoy tratando de entender: me dedico a poner ejemplos. Por ejemplo, los matemáticos llegarían con un teorema terrible... Mientras me están contando las condiciones del teorema, yo construyo algo que cumple todas las condiciones. Sabes, tienes un conjunto (una bola) –inconexo (dos bolas)–. luego, a medida que ellos añaden condiciones, en mi cabeza las bolas se vuelven de colores, les sale pelo, o lo que sea. Finalmente, definen el teorema, que es alguna tontería sobre la bola que no es cierta para mi bola verde peluda, así que digo, ¡“Falso”!

* * *

No todas las explicaciones son causales. Explicamos pruebas esbozando sus principios generales. Explicamos nuestras acciones dando nuestras razones para ellas. tom podría decir, por ejemplo, “Maté a Jerry porque se escapó con mi mujer”. Esta es su razón. Pero nosotros podríamos decir que la causa para el asesinato fue la loca envidia de tom y su incapacidad para controlarla. la mayoría de las explicaciones de la vida cotidiana sí parecen depender de relaciones causales. Enlazamos cadenas a partir de eslabones básicos. transferimos cadenas a partir de modelos similares. Cuando explicamos un suceso, tanto si ofrecemos una teoría científica totalmente montada como si es una abducción cotidiana informal, el resultado es una poderosa restricción sobre la inducción. Reduce el número de posibles inducciones; también facilita la generalización en base a una única observación de un suceso. las restricciones de una teoría a menudo hacen

caso omiso del puro proceso inductivo: ignoramos los contraejemplos a la teoría. El científico y aforista alemán del siglo dieciocho Georg lichtenberg señaló una vez: “Uno no debería tomar nota de las experiencias contradictorias hasta que haya las suficientes como para construir un nuevo sistema que merezca la pena.” El biólogo molecular James Watson ha hecho también la observación de que ningún buen modelo científico da nunca cuenta de todos los datos porque algunos de los datos tienen que ser falsos. la mayoría de nosotros parecemos seguir esta prescripción metodológica. Incluso los niños que están intentando resolver un problema ignoran los contraejemplos a las hipótesis que tienen en ese momento. Este desprecio de los datos implica que la pura intuición tiene un papel limitado en el desarrollo de las explicaciones. Una explicación no solo aumenta la información de las observaciones: puede eliminar observaciones como erróneas.

CAPÍTULO 15

El balance de probabilidades

Va usted a su médico a hacerse un chequeo anual, y la doctora le prescribe varias pruebas. Más tarde, le dice que ha dado positivo para la enfermedad de Pel Ebstein, y que los sujetos con esa enfermedad dan positivo en 99 ocasiones de cada 100. la enfermedad es seria, pero siempre que se trate sin dilación usted estará bien. Es natural estar preocupado al recibir esas noticias, pero si infirió que probablemente tiene la enfermedad, sucumbió a una falacia sutil. En su debido momento explicaré en qué consiste la falacia. Es un ejemplo típico de lo que puede fallar en nuestro razonamiento sobre probabilidades. Este razonamiento es difícil, especialmente porque el cálculo de probabilidades es una invención intelectual, y no es innato en nosotros. Aristóteles, que inventó la lógica, no nos dijo nada sobre cómo debíamos razonar sobre probabilidades. tenía poco que decir sobre ellas aparte del hecho de que los sucesos probables con frecuencia se producen. El cálculo de probabilidades, como todo jugador de póquer novato sabe, no fue inventado hasta 1654, cuando el Caballero de Méré escribió a Pascal en relación con un juego de apuestas, y Pascal a su vez consultó con Fermat. Por supuesto, la gente razonaba sobre probabilidades antes de llegar a 1650, pero nadie sabía cómo debía razonar sobre ellas. En otras palabras, el cálculo de probabilidades nos permite deducir las respuestas correctas sobre ciertas probabilidades. Proporciona, como dicen los psicólogos, una teoría “normativa”. Extrañamente, sin embargo, el cálculo no nos dice qué significan las probabilidades. Podemos interpretarlas como frecuencias de sucesos, medidas de la creencia subjetiva o tendencias objetivas de instrumentos físicos, como por ejemplo las ruletas. los teóricos debaten sin fin cuál es la interpretación adecuada.

Pocos dominamos el cálculo de probabilidades o formamos una opinión estable sobre cómo interpretarlo, pero eso no nos impide razonar sobre probabilidades. Una vía es inductiva, aproximadamente como yo imagino que razonaba Aristóteles. Sabemos que si Paolo fue a buscar el coche estará de vuelta en cinco minutos. también sabemos que fue a buscar el coche. Pero después de un cuarto

de hora no ha regresado. Así que, ¿cuál de estos hechos es más probable que se haya producido?

Se metió en una zona de calles de un solo sentido y ha tenido que salir de la ciudad.

Tuvo que salir de la ciudad.

La mayoría de la gente elige la primera opción (cuando aparece entre otra serie de opciones). Es una inducción, porque la elección es falible. Ciertamente, la primera opción es solo un caso especial de la segunda, y por tanto no puede ser la más probable de las dos. Por supuesto, la yuxtaposición de las dos opciones sugiere que en el segundo caso Paolo tuvo que salir de la ciudad por alguna otra razón. Pero, ¿por qué debería eso hacer menos probable el hecho? En este tipo de opiniones, utilizamos las pruebas y el conocimiento general para hacer inferencias sobre lo probable de un suceso. Si alguien nos presiona – quizá un inoportuno psicólogo– estamos incluso preparados para poner números a nuestras estimaciones: a partir de nuestro conocimiento de las ciudades italianas inducimos que la probabilidad de que Paolo se metiera en una zona de calles de un solo sentido es de una sobre tres. los psicólogos hacen inducciones similares sobre sus experimentos cuando infieren que tienen un efecto real debido a lo pequeña que es la probabilidad de que se hayan producido por azar.

Kahneman y tversky llevaron a cabo una serie magistral de experimentos que examinaban las inducciones intuitivas de las probabilidades. los resultados dieron la vuelta a la concepción clásica de la racionalidad humana –la opinión de que el cálculo de probabilidades era solo sentido común–. Sus estudios mostraron, por ejemplo, que si el conocimiento hace que tengamos acceso a una posibilidad, entonces tendemos a pensar que esa posibilidad tiene una probabilidad más alta que los sucesos a los que tenemos menos acceso. Alguien nos pregunta:

¿Qué es más probable en los Estados Unidos: morir en un accidente de coche, por un derrame cerebral o por cáncer de estómago?

Tendemos a decir, “accidente de coche”. En realidad, la causa más frecuente de muerte es el derrame cerebral, seguida del cáncer de estómago, seguida del accidente de coche. tenemos más acceso a las muertes por accidente de coche: son un tema habitual en los medios, mientras que las muertes de la mayoría de las personas no se registran y se tiende a morir de enfermedades. la facilidad de acceso es un heurístico que tendemos a utilizar en opiniones intuitivas inconscientes sobre probabilidades.

Cuando se nos pregunta por la probabilidad de que un sujeto sea miembro de un determinado conjunto o categoría, a menudo utilizamos en qué medida él o ella parecen ser representativos de esa categoría. Este heurístico es sensato, pero puede llevarnos a error si no comprobamos nuestras intuiciones. Un ejemplo famoso de los estudios de tversky y Kahneman se refiere a una mujer llamada linda:

Linda tiene 31 años, soltera, franca y muy brillante. Se especializó en filosofía. Como estudiante estaba muy interesada por los temas de discriminación y justicia social y también participó en manifestaciones antinucleares.

A partir de la descripción, los participantes en muchos experimentos dieron un mayor índice de probabilidad a que linda sea una cajera de banco feminista que a que sea una cajera de banco. Su descripción es más representativa de lo primero que de lo segundo (como mostraron índices independientes). Pero el conjunto de cajeras de banco incluye a las cajeras de banco feministas, y por tanto sus índices de probabilidades violan un principio básico de la probabilidad. Nuestra intuición está respondiendo a la pregunta equivocada, y nuestro poder deductivo está pasando por alto el error.

Sí que hacemos deducciones sobre probabilidades. El premio está detrás de una de las tres puertas. Usted aventura que está detrás de la puerta de la izquierda. Yo sé su verdadera localización, y abro la puerta de la derecha para mostrarle que no

está allí. ¿Debería cambiar y elegir la puerta central o mantenerse satisfecho? Usted deduce que el premio está o detrás de la puerta de la izquierda o detrás de la puerta central, y por tanto las apuestas están igualadas. Y por tanto usted sigue con su opción. la mayoría de las personas lo hace. Pero se equivocan, y más tarde verá por qué. En este tipo de razonamiento deductivo, hay varias posibilidades mutuamente excluyentes, y se utilizan para deducir probabilidades.

Los psicólogos quieren entender cómo hacen esas inferencias sujetos novatos. Por novato me refiero a quienes no han llegado a dominar la pertinente pieza de la tecnología intelectual, el cálculo de probabilidades. Se incluyen entre ellos desde los jueces de la Corte Británica de Apelaciones, que rechazaron su utilización en las pruebas judiciales, a su amigo del bar que opta por no cambiar las puertas en el problema anterior.

Algunos psicólogos rechazan los resultados de Kahneman y tversky y sostienen que gran parte de nuestro razonamiento se basa en procesos mentales, que, sin nosotros saberlo, plasman principios del cálculo de probabilidades. Estos principios permanecen dormidos a menos que la información se presente en forma de frecuencias de aparición natural. Según esta versión, cada persona desde Aristóteles hasta quienes pertenecen a culturas que solo cuentan “uno”, “dos”, “muchos”, está equipada con una encarnación inconsciente de partes del cálculo de probabilidades. Sin embargo, las capacidades de los sujetos novatos se pueden describir mejor en función de principios simples que gobiernan la inferencia de probabilidades a partir de modelos mentales. Esta hipótesis no niega el estatus normativo del cálculo para inferencias sobre probabilidades. Somos racionales al inferir probabilidades hasta el punto de que acatamos sus principios, aunque, de vez en cuando, la vida o los psicólogos nos atraen hasta inferencias irracionales.

Mi plan es explicar cómo los modelos subyacen en las deducciones sobre probabilidades, y decir algo sobre el cálculo. Aumentaré la complejidad de las inferencias hasta llegar a las que están fuera de los límites de nuestra competencia, y explicaré por qué nos resultan tan difíciles. Algunas personas sufren de náuseas existenciales tan pronto como ven números, así que si usted es una de ellas, por favor, aparte su vista (y ni siquiera examine las dos tablas).

* * *

En capítulos anteriores he sostenido que el razonamiento se basa en modelos mentales de las posibilidades, y que cada modelo mental representa una posibilidad diferente. Por ello, dada una representación de lo probables que son esas posibilidades, los modelos proporcionan una forma de razonar sobre probabilidades. Si sabemos, pongamos por caso, que el premio está detrás de una de las tres puertas, entonces, a menos que exista información de lo contrario, asumimos que las opciones de que esté detrás de cada una son iguales. Esta noción de que las posibilidades alternativas tienen iguales probabilidades tiene una larga historia en el desarrollo del cálculo de probabilidades, y se conoce a veces como el principio de “indiferencia”. Existen pruebas sólidas de que seguimos este principio, aunque asignemos las probabilidades iguales a modelos de las posibilidades. Dada una proposición, como por ejemplo: Dentro de la caja hay una canica roja o una canica verde, o ambas, los participantes en un experimento estimaban la probabilidad de que las dos canicas estuvieran dentro de la caja como de aproximadamente un tercio. Esta estimación se sigue de asumir que pensaban en las tres posibilidades como igualmente probables: canica roja dentro de la caja, canica verde dentro de la caja, canicas roja y verde dentro de la caja. Calculaban entonces la proporción de posibilidades en que se producía la situación crítica. El experimento examinaba una variedad de problemas de diferentes tipos y corroboraba la asunción de equiprobabilidad.

Estos resultados podrían estar abiertos a otras explicaciones, por tanto, ¿cómo pueden los psicólogos mostrar que utilizamos modelos mentales? la respuesta es que la teoría predice la aparición de sesgos sistemáticos, debido a que los modelos mentales representan lo que es verdadero, pero no lo que es falso. Este principio de verdad, que discutí anteriormente (en el Capítulo 8), se aplica a dos niveles: construimos modelos de las posibilidades verdaderas, y cada modelo representa solo aquellas proposiciones que son verdaderas dentro de él. Como ejemplo, consideremos este problema:

Hay una caja que tiene dentro al menos una canica roja, o bien una canica verde y una canica azul, pero no las tres canicas. ¿Cuál es la probabilidad de que dentro de la caja haya una canica roja y una canica azul?

Como predicen los modelos mentales de la proposición, tendemos a pensar en dos posibilidades. En una posibilidad hay una canica roja y en la otra hay una canica verde y una canica azul. Por tanto, es imposible para las canicas en cuestión –la roja y la azul– aparecer juntas. la probabilidad de que ambas estén dentro de la caja es por tanto cero. Nuestro razonamiento imperfecto es un perfecto ejemplo del uso de modelos mentales. Vamos a rehacer las etapas inferenciales, pero esta vez vamos a conseguir la respuesta correcta.

La premisa significa que cuando es verdadero que dentro de la caja hay una canica roja, es falso que haya una canica verde y una canica azul, y viceversa –esta es la fuerza del "o bien"–. la falsedad de la conjunción de una canica azul y una verde puede producirse de tres formas: verde sin azul, azul sin verde y ni verde ni azul. Cada una de estas posibilidades puede coincidir con la canica roja. la premisa significa también que cuando es verdadero que dentro de la caja hay una canica verde y una canica azul, es falso que haya una canica roja. Y por tanto hay cuatro posibilidades distintas respecto a lo que hay dentro de la caja:

Roja Verde

Roja Azul

Roja

Verde Azul

Si cada una de estas posibilidades es equiprobable, entonces la probabilidad de que dentro de la caja haya una canica roja y una azul no es cero, sino un cuarto.

Un experimento examinó varias inferencias de este tipo, que la teoría predice que deberían dar lugar a ilusiones, y las comparó con inferencias a partir de

problemas de control, que se correspondían con ellas, en las que los modelos mentales dan lugar a las respuestas correctas. He aquí un ejemplo de un problema de control:

Hay una caja que tiene dentro una canica roja o bien una canica verde, y una canica azul, pero no las tres canicas. ¿Cuál es la probabilidad de que haya una canica roja y una canica verde dentro de la caja?

La premisa sugiere de nuevo que hay dos posibilidades, dependiendo de si dentro de la caja hay una canica roja o una canica verde, pero en ambas posibilidades hay una canica azul. Por tanto, la teoría predice que es imposible que ambas canicas: roja y verde estén dentro de la caja. Y esta vez la respuesta predicha es correcta. Solo hay esas dos posibilidades. los participantes sucumbieron a las ilusiones y lo hicieron mucho mejor con los problemas de control. tres cuartas partes de los participantes estimaron una probabilidad de cero, o algo cercano, para la proposición en el ejemplo ilusorio anterior, y solo un par de participantes hicieron la estimación no sesgada. En cambio, la mayoría de los participantes dieron una respuesta correcta al problema de control. Estos resultados eran representativos del experimento como conjunto.

En general, podemos inferir la probabilidad de un suceso representando cada una de las posibilidades mutuamente excluyentes en un modelo independiente. Asumimos, a menos que tengamos pruebas en contra, que cada posibilidad es equiprobable. Computamos la probabilidad de una proposición calculando la proporción de las posibilidades en que aparece. Pero cometemos errores, porque tendemos a no representar lo que es falso.

* * *

Phil tiene dos hijos. Uno es un chico. ¿Cuál es la probabilidad de que el otro sea un chico? la mayoría de nosotros razonamos de esta forma: hay dos posibilidades para el segundo hijo de Phil: un chico o una chica. Ergo, la

probabilidad de que el segundo hijo sea un chico es un medio. Verosímil, pero erróneo. la forma correcta de proceder es enumerar todas las posibilidades:

Hijo nacido en primer lugar	Hijo nacido en segundo lugar
Chica	Chica
Chica	Chico
Chico	Chica
Chico	Chico

El hecho de que Phil tenga un hijo elimina la primera de las posibilidades. Quedan tres posibilidades, y solo en una de ellas ambos hijos son chicos. la respuesta correcta por tanto es un tercio. Si yo hubiera dicho que el primogénito de Phil era un chico, entonces la probabilidad de que su segundo hijo fuera un chico hubiera sido realmente un medio. Pero enmarqué la pregunta de forma que no hubiera manera de identificar cuál de los hijos de Phil se sabía que era chico.

El problema precedente gira sobre una probabilidad condicional. Podemos hacer frente a otras probabilidades condicionales. Por ejemplo, hay dos bolsas, y una de ellas contiene una moneda y la otra está vacía. Elijo una de las bolsas al azar. Si contiene la moneda, giro la moneda, y si cae cara, usted gana. Pero, en cualquier otro caso, pierde. Suponiendo que todo se hace sin trampa, ¿cuál es la probabilidad de que usted gane? Debe pensarlo por sí mismo:

Es 50:50 de que Phil elija la bolsa con la moneda, y luego es 50:50 de que caiga cara. Por tanto la probabilidad para la bolsa con la moneda es un medio, y dada esa bolsa, la probabilidad de que la moneda caiga cara es un medio, y por tanto la probabilidad de que yo gane es de un medio de un medio, lo que supone un cuarto.

Usted multiplica las dos probabilidades y obtiene la respuesta correcta. Para probarla, he aquí el conjunto completo de posibilidades con las opciones de aparición, que he expresado como frecuencias idealizadas en cien repeticiones del juego:

Phil coge la bolsa con la moneda: 50. Cae cara: 25.

Cae cruz: 25.

Phil coge la bolsa vacía: 50.

Por tanto, las opciones de que usted gane son ciertamente veinticinco veces de cada 100, o un cuarto. Un análisis correcto depende de su conocimiento de que cuando deben producirse dos sucesos que son independientes uno de otro, su probabilidad conjunta es el producto de sus probabilidades individuales. la mayoría de nosotros retiene al menos este conocimiento de sus días escolares, o bien puede imaginar los modelos apropiados –parecidos a los de más arriba– etiquetados con las probabilidades numéricas adecuadas.

Sin embargo, las probabilidades condicionales están en el límite de nuestra competencia. lo que no funciona en el problema sobre los hijos de Phil es que lo comprendemos mal. No nos damos cuenta de que es una cuestión de probabilidad condicional: dado que un hijo es chico, ¿cuál es la probabilidad de que el otro hijo sea un chico? la frase “dado que...” es una clave que se nos escapa. Como consecuencia, no pensamos sobre el conjunto completo de posibilidades, y en lugar de eso consideramos que la pregunta se refiere a una probabilidad absoluta simple: ¿cuál es la probabilidad de que un hijo sea chico?

Nuestra tendencia a imaginar un conjunto de cuatro posibilidades diferentes no es mayor para inferir una probabilidad que para hacer una deducción normal.

Son demasiadas para que las tengamos en mente a la vez, y por eso tendemos a pensar en ellas de una en una. Como ejemplo que viene al caso, volvamos al problema del inicio del capítulo:

Si los sujetos tienen la enfermedad de Pel Ebstein entonces dan positivo en 99 de cada 100 ocasiones. Usted da positivo. ¿Cuál es la probabilidad de que tenga la enfermedad?

La mayoría de los participantes en un experimento respondieron con una probabilidad alta a cuestiones de este tipo. Un experto en el cálculo de probabilidad sería reacio a responder. Para ver por qué, puede usted consultar la tabla 15.1, que establece las posibilidades y sus frecuencias. la moraleja de la tabla es que una probabilidad condicional como por ejemplo:

Tabla 15.1. El riesgo de tener la enfermedad de Pel Ebstein

He aquí las posibilidades y sus frecuencias para la probabilidad condicional de que una persona con la enfermedad de Pel Ebstein dé positivo 99 veces de cada 100:

Frecuencias

Tiene la enfermedad de Pel Ebstein	Da positivo	99
Tiene la enfermedad de Pel Ebstein	No da positivo	1

No tiene la enfermedad de Pel Ebstein	Da positivo	No se sabe
No tiene la enfermedad de Pel Ebstein	No da positivo	No se sabe

Quiere usted saber la probabilidad condicional de tener la enfermedad dado que da positivo. Por tanto, necesita saber las frecuencias de los casos en las filas primera y tercera en las que da positivo, y luego en qué proporción de ellas tiene usted la enfermedad (la primera fila). No tiene información sobre la frecuencia con la que los sujetos dan positivo y no tienen la enfermedad (la tercera fila). Supongamos que las frecuencias completas son las siguientes:

Frecuencias

Tiene la enfermedad de Pel Ebstein	Da positivo	99
Tiene la enfermedad de Pel Ebstein	No da positivo	1
No tiene la enfermedad de Pel Ebstein	Da positivo	1.000.000
No tiene la enfermedad de Pel Ebstein	No da positivo	2.000.000

En este caso, aunque usted dio positivo, el riesgo de que padezca la enfermedad es pequeño: 99/1.000.099, es decir, apenas una posibilidad entre diez mil.

Esta forma de inferir la probabilidad condicional de A dado B se conoce como el método de subconjuntos, porque depende de la relación:

La probabilidad de la conjunción de A y B (el subconjunto)

La probabilidad de B (el conjunto)

Una vez que usted conoce la probabilidad para cada conjunción de sucesos, como en la tabla anterior, puede calcular cualquier probabilidad condicional utilizando el método de subconjuntos. Sabe todo lo que hay que saber sobre las probabilidades en ese campo.

Si usted tiene la enfermedad entonces da positivo en 99 de cada 100 ocasiones

no implica la probabilidad condicional conversa:

Si usted da positivo entonces usted tiene la enfermedad en 99 de cada 100 ocasiones.

Una probabilidad condicional no le dice a usted casi nada sobre la otra probabilidad condicional a no ser que tenga información adicional (véase la tabla 15.2). Sin embargo, los participantes en el experimento no rehusaron ni una sola vez hacer toda una serie de inferencias de ese tipo tanto sobre probabilidades altas como sobre probabilidades bajas. ¿Por qué? No asumieron por sistema que las dos probabilidades condicionales fueran iguales. Pero, en el ejemplo médico no conocían ninguna causa alternativa para dar positivo. El experimento incluía también problemas en los que había alternativas obvias, como por ejemplo:

Si un objeto se mete en agua hirviendo, entonces las opciones de que se caliente son de 99 sobre 100. Un objeto está caliente. ¿Cuál es la probabilidad de que se haya metido en agua hirviendo?

Los participantes podrían pensar en otras muchas posibilidades que podrían hacer que el objeto estuviera caliente, p. ej., había estado al sol, se había metido en un horno, se había sometido a una intensa fricción. No se mostraron reacios a inferir una probabilidad, pero era más baja que la probabilidad establecida en el problema. Después del experimento decían que habían pensado sobre las formas alternativas mediante las que podría haberse calentado un objeto y las habían tenido en cuenta al hacer su inferencia. Pero inferían también probabilidades condicionales conversas sin tener el conocimiento pertinente para ejemplos similares al caso de la enfermedad de Pel Ebstein.

La teoría de modelos explica su comportamiento. Postula que tendemos a pensar sobre las posibilidades de una en una. Dada una probabilidad condicional, como por ejemplo, si una persona tiene fiebre de Pel Ebstein entonces da positivo 99 veces de cada 100, tendemos a construir, uno por uno, los dos modelos mentales siguientes:

Tiene la enfermedad de Pel Ebstein	Da positivo 99
...	1

La elipsis representa una posibilidad alternativa que no hacemos explícita. Cuando se nos dice que una persona dio positivo, y se nos pregunta por la probabilidad condicional de que la persona tenga la enfermedad, asumimos que se aplican los mismos modelos, e inferimos el mismo valor: 99 de cada 100. De manera similar, si en lugar de eso se nos dijera que la persona no dio positivo, asumimos que la posibilidad corresponde al modelo implícito. Y por ello ahora inferimos una probabilidad baja de que la persona tenga la enfermedad. Cuando pensamos en otras posibilidades, como sucede en el ejemplo sobre el agua

hirviendo, también construimos modelos para esas posibilidades. luego estimamos una probabilidad más baja para la probabilidad condicional conversa.

Consideremos, de nuevo, el problema sobre los dos hijos de Phil. Dado que un hijo es chico, necesitamos evaluar la probabilidad condicional de que el otro hijo sea chico. Para calcular la respuesta correcta, necesitamos pensar en las tres posibilidades en las que al menos un hijo es chico:

Chica Chico

Chico Chica

Chico Chico

Encontramos entonces el subconjunto de posibilidades en el que el otro hijo es chico. Solo hay una posibilidad de ese tipo. Calculamos la proporción del subconjunto respecto al conjunto como un todo: 1/3. Y ese resultado es la probabilidad condicional requerida. la mayoría de nosotros somos capaces de seguir este método de subconjuntos para evaluar una probabilidad condicional a partir de modelos explícitos. Varios estudios han mostrado que los participantes novatos son capaces de inferir probabilidades condicionales por este camino con tal de que puedan llevar a cabo el proceso paso a paso y la aritmética sea fácil.

* * *

Hay dos sucesos, y cada uno de ellos tiene una probabilidad de ocurrir de un medio. ¿Cuál es la probabilidad de que se produzcan ambos? Aquellos que no están familiarizados con el cálculo de probabilidades tienden a decir: un cuarto.

Multiplican entre sí dos probabilidades para inferir la probabilidad de una conjunción de dos sucesos. En el ejemplo anterior de la bolsa y la moneda, este procedimiento es correcto. Pero, en este ejemplo, es erróneo. Podemos ver inmediatamente nuestro error si nos enteramos de que un suceso era el lanzamiento de una moneda que sale cara y el otro suceso era el mismo lanzamiento de la moneda saliendo cruz. No pueden producirse los dos sucesos, y por tanto la probabilidad de su conjunción es cero. la descripción era vaga con el fin de atraernos al error. No nos decía nada sobre la relación entre los dos sucesos. Podían ser dependientes uno de otro como en las caras y las cruces. Pero también podían ser independientes como en los lanzamientos de dos monedas diferentes.

Permítame hacer una digresión en este asunto de la dependencia. El grado de dependencia entre dos sucesos, como el que yo me acuerde de coger un paraguas y que más tarde llueva ese día, se refleja en la probabilidad condicional de que dado un suceso se produzca el otro. Supongamos que el "índice base" de lluvia en un día, es decir, la probabilidad de que llueva, es uno sobre diez. Y supongamos que la probabilidad condicional de que llueva en el caso de que yo recuerde coger un paraguas es también de uno sobre diez. Se sigue que el que llueva o no no se ve afectado por el hecho de que yo recuerde coger un paraguas: los dos sucesos son independientes. Este caso contrasta con un pronóstico del tiempo. la probabilidad de que llueva es, pongamos por caso, nueve sobre diez en el caso de que el pronóstico sea de lluvia, mientras el índice base de lluvia es solo uno sobre diez. Esta diferencia muestra que la lluvia y el pronóstico de lluvia no son sucesos independientes. Ciertamente, el pronóstico es útil, porque cuando predice lluvia, la probabilidad de lluvia es más alta que el índice base. En general, dos sucesos son independientes si la probabilidad de un suceso es la misma que su probabilidad condicional en el caso de que se produzca el otro suceso.

Si estamos jugando a la ruleta y observamos una secuencia de doce rojos uno tras otro, podríamos pensar que en el siguiente giro de la rueda debe ser más probable conseguir negro con el fin de asegurar que el reparto entre los dos colores sea aproximadamente 50:50. Esta forma de pensar se conoce como la "falacia del jugador". Dada una ruleta sin trampa, cada resultado es independiente de los otros, y las opciones para el rojo y negro son iguales en cada giro. Por tanto, una secuencia previa de rojos no tiene efecto sobre la probabilidad del rojo en el siguiente resultado. la implicación real de una secuencia de doce rojos se ilustra en lo que sucedió cuando se legalizaron por

primera vez los casinos en el Reino Unido en los años sesenta. Un estadístico americano me dijo que en una visita que hizo entonces al Reino Unido observó que los casinos no emparejaban los rotores de las ruletas –las partes que giran– con bases diferentes de una sesión a otra. Por tanto, él hacía pequeñas apuestas durante un largo rato y guardaba un registro de las frecuencias con que salían los diferentes números. Después de un día o dos de juego, algunas veces descubría que una rueda tenía un sesgo en el que algunos números salían más a menudo de lo que uno esperaría de la casualidad. llevó a cabo una prueba estadística para descubrir si el sesgo era significativo. Si era significativo, posteriormente apostaría a esos números, y ganaría una gran cantidad de dinero. Continuó utilizando este método durante unos días hasta que el mánager del casino se dio cuenta de lo que estaba pasando, e introdujo la política de emparejar los rotores y las bases de forma diferente cada día. Cuando realmente se produce una secuencia larga de sucesos improbables, la inferencia errónea es seguir la falacia del jugador. Deberíamos seguir observando la secuencia, porque los datos pueden mostrar que el proceso está sesgado.

Ahora, podemos entender cómo calcular la probabilidad de la conjunción de dos sucesos. Cuando los dos sucesos son independientes –como en mi ejemplo de la bolsa y la moneda, multiplicamos sus probabilidades para obtener la probabilidad de su conjunción–. Pero una forma segura de evaluar la probabilidad de la conjunción de dos sucesos, incluso cuando son dependientes, es multiplicar el índice base de un suceso por la probabilidad condicional de que dado ese suceso se produzca el otro. En la primera oración de la tabla 15.2 aparece descrito este principio.

Una probabilidad condicional por sí misma no nos dice casi nada sobre la probabilidad condicional conversa. Por tanto, ¿qué información necesitamos realmente para inferir esa probabilidad conversa? El cálculo de probabilidades proporciona formas alternativas de computar su valor. Si conocemos el conjunto completo de probabilidades para cada conjunción de sucesos, sabemos todo lo que hay que saber sobre las probabilidades en cuestión. En ese caso, la tabla 15.1 que aparecía anteriormente muestra cómo utilizar el método de subconjuntos. No obstante, un método alternativo depende de un teorema famoso en el cálculo de probabilidades, que toma su nombre del ministro protestante inglés el reverendo thomas Bayes, el descubrimiento del mencionado teorema por Bayes apareció en una publicación póstuma en 1764. En lugar de conocer la distribución completa de probabilidades para cada conjunción de sucesos (como en la tabla 15.1), necesitamos conocer tres probabilidades: la probabilidad condicional que es

nuestro punto de partida, p. ej., si una persona tiene la enfermedad de Pel Ebstein la probabilidad de que esa persona dé positivo es 99 sobre 100, la probabilidad de que una persona tenga la enfermedad de Pel Ebstein y la probabilidad de que una persona dé positivo. Armados con estas tres posibilidades, el teorema de Bayes, que explico en la tabla 15.2, nos permite computar la probabilidad condicional conversa. Si el teorema le suena a chino, no se preocupe. Eso nos pasa a la mayoría. No conocemos el teorema; lo encontramos difícil de entender, y, como muestran las pruebas, no lo utilizamos para inferir probabilidades condicionales.

Tabla 15.2. Teorema de Bayes y el ejemplo de Pel Ebstein

1. En general, la probabilidad de A y B = (probabilidad de A) (probabilidad de B dado A). Es también el producto de la probabilidad de B y la probabilidad de A dado B.

2. la relación para el subconjunto a partir de la tabla 15.1 es:

$$\text{probabilidad (A y B)}$$

La probabilidad de A dado B =

$$\text{probabilidad (B)}$$

3. Para la probabilidad de A y B en este caso, podemos sustituir su equivalente a partir del punto 1 anterior

$$\text{probabilidad de A dado B} = \frac{\text{probabilidad (A) (probabilidad de B dado A)}}{\text{probabilidad de (B)}}$$

Este es el teorema de Bayes, que interrelaciona una probabilidad condicional con su conversa.

4. Como ejemplo, consideremos las tres posibilidades siguientes de la tabla 15.1, en las que "≈" significa "aproximadamente igual":

A: la probabilidad de que una persona tenga la enfermedad es 100/3.000.100 ≈ 1/30.000

B: la probabilidad de que una persona dé positivo es 1.000.099/3.000.100 ≈ 1/3

B dado A: la probabilidad de una prueba positiva dada la enfermedad es 99/100.

El teorema de Bayes establece: la probabilidad de la enfermedad dada una prueba

$$\text{positiva} = \frac{(1/30.000)\,(99/100)}{1/3} \approx 1/10.000$$

Por tanto, dado que una persona dé positivo, la probabilidad de que la persona tenga la enfermedad de Pel Ebstein es baja, aproximadamente una entre diez mil.

Como expliqué al principio del capítulo, a veces hacemos inducciones intuitivas sobre probabilidades, basándonos en heurísticos simples, como la accesibilidad de conocimiento relevante. Pero también deducimos probabilidades a partir de nuestro conocimiento de las posibilidades alternativas y lo probable que resulta que se produzcan. En la vida real, los dos métodos están a menudo entremezclados. Pero, cuando un método se enfrenta al otro, el conocimiento relevante es más vívido y produce una impresión mayor. En un famoso estudio, tversky y Kahneman compararon ambos, y descubrieron que las descripciones intuitivas triunfaban sobre las probabilidades numéricas. Se les decía a los participantes, por ejemplo, que un grupo de psicólogos había pasado test de personalidad a treinta abogados y setenta ingenieros. Se les daban luego a los participantes las descripciones de cinco de esos sujetos seleccionados al azar, y se les pedía que valoraran la probabilidad de que cada uno de los sujetos fuera ingeniero. también se les dijo que se les daría una paga extra si su estimación se aproximaba a la de un grupo de expertos. Si los participantes fueran sensibles a estas frecuencias, entonces sus valoraciones deberían tener en cuenta los números respectivos de abogados e ingenieros. Pero solo lo hicieron en sus estimaciones cuando no se les dio absolutamente ninguna información sobre un sujeto. Con una descripción neutral de un sujeto tendían a valorar la probabilidad de que fuera ingeniero como 50:50. Y con una descripción vívida de una persona que parecía un ingeniero le valoraban como tal, incluso si se les había dicho que había 30 ingenieros y 70 abogados.

* * *

Nuestra educación y bagaje cultural afectan a nuestras opiniones sobre la probabilidad. la mayor diferencia se produce si en nuestra cultura existe algún conocimiento sobre cálculo de probabilidades. Como señaló una vez un filósofo, “cualquiera que hubiera jugado a los dados en tiempos de los romanos equipado con el cálculo de probabilidades habría ganado rápidamente toda la Galia”. Una feliz consecuencia de nuestra ingenua comprensión de las probabilidades, sin embargo, es que nos basta para adquirir conocimiento experto. Esta pericia, sin

embargo, depende de diversos tipos de conocimiento. Uno es el conocimiento de combinaciones y permutaciones. Por ejemplo, si Phil tiene cuatro hijos es más probable que tenga dos chicos y dos chicas, ¿no? No, no lo es. Parece verosímil, porque las oportunidades de chico y chica son iguales. los expertos saben que es necesario tener en cuenta todas las posibilidades, y que es más probable que Phil tenga tres hijos de un sexo y uno del otro sexo. El que este tipo de pericia sea resultado de la observación o del desarrollo del conocimiento conceptual es un punto discutible. otro elemento de la pericia es la capacidad para la aritmética –un factor que es fácil de pasar por alto en estudios sobre razonamiento probabilístico–. Podemos hacer cálculos aritméticos más fácilmente con frecuencias que con fracciones o porcentajes. Estos elementos nos permiten aprender el cálculo de probabilidades, y la capacidad para adquirir esta pericia reside en los sencillos principios de razonamiento que he explicado en este capítulo.

* * *

Cuando hacemos inducciones sobre probabilidades basamos nuestras opiniones en diversos heurísticos, como la accesibilidad del conocimiento relevante. Cuando hacemos deducciones sobre probabilidades representamos las diferentes posibilidades en modelos mentales, aunque como regla general no representamos lo que es falso. A menos que tengamos conocimiento en sentido contrario, asumimos que cada modelo representa una alternativa equiprobable. Inferimos la probabilidad de un suceso a partir de la proporción de modelos en que aparece, utilizando a veces etiquetas numéricas sobre los modelos para representar las probabilidades. las probabilidades condicionales las inferimos de la misma forma, considerando subconjuntos de modelos. Cualquier probabilidad que no podamos inferir por estas vías puede estar fuera de nuestro alcance. todos cometemos errores. No nos damos cuenta de que una probabilidad condicional no nos dice casi nada sobre la probabilidad condicional conversa. Ninguno de nosotros parece tener un conocimiento inconsciente de las leyes de la probabilidad. Si el doctor le dice que quienes tienen la enfermedad de Pel Ebstein dan positivo 99 de cada 100 veces, y que ha dado positivo, no hay necesidad de alarmarse. tiene usted que calcular las frecuencias relativas de todas las posibilidades. Si no puede rellenar todas las casillas, entonces cuidado

con las proposiciones sobre probabilidades condicionales.

Me queda un punto del tema sin concluir –el problema de las tres puertas–. Recordará usted que había un premio detrás de una de las tres puertas, y que usted eligió la de la izquierda. las opciones de que el premio esté detrás de esa puerta son una de tres, y las opciones de que esté detrás de una u otra de las dos puertas restantes son de dos de tres. Yo conozco la situación real del premio, y abro la puerta de la derecha para mostrarle que el premio no está tras ella. Mi acción no cambia la localización del premio o la probabilidad de que su elección sea correcta. Por tanto, las opciones de que el premio esté detrás de la puerta de la izquierda siguen siendo una de tres, y las opciones de que el premio esté detrás de la puerta central son dos de tres, porque he eliminado la posibilidad de que esté detrás de la puerta de la derecha. le ofrezco a usted la oportunidad de cambiar su elección por la de la puerta central. la mayoría de las personas rehúsan, pero usted debería cambiar a la puerta central, porque las opciones de que el premio esté detrás de ella son dos de tres. Evidentemente, puede usted utilizar el teorema de Bayes para calcular la respuesta, pero sería la misma. Si se equivoca no se alarme. Está en excelente compañía.

Parte V

Qué nos hace racionales

Lo último por lo que nos regimos es un antimodelo

Alguna falsedad que cualquiera puede desmentir

Una nada en cuya existencia nadie tiene por qué creer

Auden

CAPÍTULO 16

Los contraejemplos

¿Qué es un contraejemplo? Bien, supongamos que yo le digo:

Todo el mundo que fuma marihuana llega a tomar drogas duras.

Usted contraataca con un contraejemplo:

Eso no es cierto. Bill fumaba marihuana, pero no llegó a tomar drogas duras.

Y yo digo:

Pero, realmente él no fumaba marihuana, porque no inhalaba.

Usted dice:

Bueno, conozco a varias personas que fumaban marihuana y no llegaron a tomar drogas duras.

En este punto, me veo obligado a hacer una concesión:

Bueno, quizá no todo el mundo, pero para la mayoría de las personas es cierto.

A usted le resulta difícil rebatir esta réplica, pero también me resulta a mí difícil mantenerla. Una forma de actuar es que hagamos un recuento completo de los fumadores de marihuana, o bien estudiemos una muestra al azar y computemos algunas estadísticas y un margen de error. Tendríamos que habernos puesto de acuerdo previamente en qué consideramos como “la mayoría de los fumadores de marihuana”.

Si en lugar de mi última afirmación, yo dijera:

Algunos fumadores de marihuana llegan a tomar drogas duras

entonces su única refutación posible sería:

Ningún fumador de marihuana llega a tomar drogas duras.

Y a partir de ese momento usted debería comprobar a cada fumador de marihuana, mientras que yo solo necesito encontrar un ejemplo o dos. Este ir y venir es como un juego, y algunos filósofos han analizado el significado de cuantificadores, como por ejemplo “todos” y “algunos”, en función de estos juegos. Pero lo que muestra el juego es que usted y yo entendemos la noción de contraejemplo, al menos para afirmaciones sobre conjuntos de individuos.

Los contraejemplos también son cruciales para el razonamiento. Una inferencia válida tiene una conclusión que debe ser verdadera si sus premisas son

verdaderas. Un antimodelo de una proposición representa una posibilidad en la que la proposición es falsa. Y por tanto un contraejemplo es un modelo de las premisas que es un antimodelo de la conclusión. Establece que la conclusión es falsa en al menos una posibilidad consistente con las premisas, y por tanto la conclusión no es una inferencia válida a partir de las premisas. Sugerí al comienzo del libro que una piedra angular de la racionalidad humana –en la medida en que los seres humanos son racionales– es que captan la fuerza de los contraejemplos. Es hora de sacar provecho de esta afirmación.

Un contraejemplo para una proposición general muestra que la proposición es falsa. Sin embargo, a menudo no buscamos contraejemplos. En el Capítulo 1 describí un famoso paradigma experimental, la tarea de selección de Wason, en la que los participantes tenían que seleccionar pruebas que mostraran si una afirmación general es verdadera o falsa. Recordará que el experimentador presenta una afirmación, como por ejemplo, “Si una carta tiene una ‘A’ en una cara entonces tiene un ‘2’ en la otra cara” e invita a los participantes a seleccionar en una serie de cuatro (A, B, 2 y 3) aquellas cartas que es necesario volver para descubrir si la afirmación es verdadera o falsa respecto a las cuatro cartas. Los participantes saben que cada carta tiene en una cara una letra y en la otra un número. Con frecuencia no seleccionan la carta que tiene un “3” en una cara, pero es un contraejemplo potencial, porque una “A” en su otra cara mostraría que la afirmación general es falsa. El principio de verdad les lleva a centrarse en una posibilidad verdadera, las cartas que tienen una A y un 2. Sin embargo, cualquier manipulación experimental que les facilita el pensar en contraejemplos parece mejorar su comportamiento en la tarea. Por tanto, somos capaces de captar la fuerza de los contraejemplos si las condiciones de un experimento nos capacitan para superar el principio de verdad. Muy a nuestro pesar, no existe consenso sobre la tarea de selección. Algunos psicólogos sostienen que no consigue evaluar el razonamiento, y una revista psicológica ha rechazado publicar ningún otro artículo sobre ella, debido a su controvertido estatus.

Otra tarea que diseñó Wason también sugiere que tendemos a no buscar contraejemplos. El experimentador nos muestra una secuencia de números, “2 4 6” y dice que son un ejemplo de una regla general que tiene en mente. Nuestra tarea es descubrir cuál es esa regla general, y procedemos como sigue. Construimos una secuencia de tres números, y el experimentador nos dice si es o no un ejemplo de su regla. Escribimos, pongamos por caso, “6 8 10”, y él dice, “sí, ese es un ejemplo de mi regla”. Construimos otros tres números, y el

experimentador responde a ellos. Y continuamos de esta forma hasta que hemos identificado la regla. La tarea no es fácil. Cuando pensamos en una hipótesis sobre la regla, como por ejemplo “los números aumentan de dos en dos”, tendemos a construir una secuencia que le sirva de ejemplo, como: 3, 5, 7. Pero es difícil resolver el problema de esta forma, porque no nos damos cuenta de que la regla del experimentador es muy general. En realidad es: tres números cualquiera de magnitud creciente. Es probable que la mayoría de nuestras hipótesis sean menos generales. Por tanto, las comprobaciones positivas de nuestra hipótesis no tienen utilidad, porque cada uno de sus ejemplos es también un ejemplo de la regla del experimentador. Y por tanto nunca podemos saber que nuestras hipótesis son erróneas. En consecuencia, algunos participantes no consiguen resolver el problema ni siquiera después de trabajar en él durante una hora. Si nuestra hipótesis es, “números que aumentan de dos en dos”, entonces una forma mejor de actuar es construir una secuencia que sea un contraejemplo para ella, como por ejemplo: 1, 2, 3. El experimentador nos dice que esta secuencia es un ejemplo de su regla, y por tanto sabemos que nuestra hipótesis es errónea: el contraejemplo de nuestra regla es un ejemplo de la suya.

¿Por qué comprobamos ejemplos positivos de nuestras hipótesis? Quizá tenemos una tendencia a confirmar nuestras hipótesis, pero una explicación más simple y verosímil es el principio de verdad. Nos centramos en ejemplos verdaderos de nuestras hipótesis, que son más fáciles de imaginar que los ejemplos falsos.

En su formulación original, la teoría de modelos postulaba que quienes razonan construyen modelos de las premisas, extraen una conclusión, y luego, dado un mínimo de competencia, buscan contraejemplos para la conclusión. Pero, ¿lo hacen? En realidad, con frecuencia construyen un modelo, extraen de él una conclusión, y eso es todo. Cuesta trabajo superar el principio de verdad, y parece que hay quien no tiene ganas o es incapaz de buscar contraejemplos. Este capítulo necesita por tanto responder a tres preguntas: ¿Reconocemos la fuerza de los contraejemplos? ¿En qué circunstancias, de existir alguna, los buscamos? Y ¿explica la teoría de modelos nuestro comportamiento?

* * *

Un problema en el uso de los contraejemplos es darse cuenta de su necesidad. Cuando valoramos la verdad de proposiciones generales no siempre captamos su importancia. De la misma manera, cuando extraemos nuestras propias conclusiones a partir de las premisas, tendemos a centrarnos en un único modelo respecto a ellas. Construir ese modelo, y formular una conclusión, es tan preocupante que podemos no buscar contraejemplos. No nos queda capacidad para pensar en ellos. Una forma de reducir la carga de nuestra capacidad mental, y nuestra memoria de trabajo, es darnos una conclusión que tengamos que evaluar, en lugar de pedirnos que formulemos una. Este procedimiento puede crear también una actitud más crítica, porque tenemos que enfrentarnos con la conclusión de otro en lugar de con la nuestra.

Otro problema con los contraejemplos es determinar qué se considera como tal. Supongamos que extraemos una conclusión que es una disyunción excluyente:

O está Pat y está Viv, o bien está Eve.

¿Qué constituye un contraejemplo para esta conclusión? Cualquier antimodelo en el que la conclusión es falsa pero las premisas son verdaderas, pero no tenemos un acceso inmediato a esos modelos. Tenemos que construirlos. Supongamos que no está Pat. Esto refutaría la conjunción inicial en la conclusión, pero si está Eve, la conclusión todavía es verdadera. Añadamos la suposición posterior de que no está Eve. Por tanto, no está Pat y no está Eve. Ese sería un contraejemplo. ¿Hay otros? Podemos imaginar la posibilidad en la que no está Viv y no está Eve. Incluso podemos imaginar la posibilidad en la que no están ninguna de las mujeres. Pero puede que no pensemos en una posibilidad final en la que ambas cláusulas de la disyunción sean verdaderas: están las tres mujeres. Sin embargo, también es un contraejemplo, si la disyunción es excluyente.

Experimentos en los que los participantes tienen que enumerar los antimodelos de proposiciones muestran que utilizan distintas formas de hacerlo. Pero su actuación sigue estando restringida por el principio de verdad, porque primero piensan en lo que es verdadero, y luego tratan de refutarlo. Por consiguiente, pueden hacer frente a proposiciones que contienen solamente una única

conectiva; aunque puede que no consigan enumerar todas las posibilidades en las que una conjunción es falsa. Sin embargo, como predice el uso de modelos mentales, su actuación con proposiciones que contienen dos o tres conectivas es con frecuencia nefasta. Cuando los participantes tienen que enumerar separadamente lo que es posible y lo que es imposible dada una proposición, es decir, sus modelos y antimodelos, casi nunca construyen dos listas que juntas agoten todas las posibilidades relevantes. La única forma segura de construir contraejemplos se describe en la Tabla 16.1, pero solo es probable que la utilicen quienes tengan una formación en lógica. Por tanto, cuando imaginamos contraejemplos, corremos el riesgo de equivocarnos.

Tabla 16.1. Cómo crear un contraejemplo para un aserto

Ejemplo: O está Pat y está Viv, o bien está Eve.

1. Construyamos todas las posibilidades en las que el aserto como un todo es verdadero. El principal conectivo es “o bien”, que podemos tratar como una disyunción excluyente, y por tanto combinar las posibilidades verdaderas para una cláusula con las falsas para la otra cláusula:

Pat Viv Eve no [Primera cláusula verdadera y segunda cláusula falsa]

Pat Viv no Eve [Primera cláusula falsa y segunda cláusula verdadera]

Pat no Viv Eve [Primera cláusula falsa y segunda cláusula

verdadera]

Pat no Viv no Eve [Primera cláusula falsa y segunda cláusula verdadera]

donde “Pat” representa “está Pat”, y “no” niega esta proposición.

2. Eliminemos este conjunto del conjunto de ocho posibles combinaciones de está Pat, está Viv, está Eve, y sus negaciones. Lo que queda es el conjunto de antimodelos del aserto original, es decir, el conjunto de posibilidades en las que el aserto es falso:

Pat Viv Eve

Pat Viv no Eve no

Pat no Viv Eve no

Pat no Viv no Eve no

Estos son los contraejemplos posibles para el aserto original. Expresado con palabras, están las tres personas, solo está Pat, solo está Viv, o no está ninguna de las mujeres.

▪

Otra forma de construir contraejemplos es pensar en la negación de una oración, y luego tratar de imaginar las posibilidades compatibles con ella. Pero no es fácil pensar en la negación correcta de muchas oraciones. Consideremos una proposición como:

Están todas las mujeres.

Solo con que no esté una mujer la proposición es falsa, y por tanto su negación adecuada es:

No están todas las mujeres.

Como muestra este ejemplo, una negación adecuada contradice la proposición que niega. Es decir, una de las dos proposiciones debe ser verdadera y una de ellas debe ser falsa. Muchos de nosotros, sin embargo, suponemos que la introducción de cualquier negación en la proposición original la niega. Por tanto, podemos pensar que la negación de la proposición es: no está ninguna de las mujeres. Pero esta proposición no contradice la proposición original, porque tanto ella como la original podrían ser falsas –como en el caso de que estén algunas mujeres pero no todas.

Por tanto, ¿cuál es la forma correcta de negar una proposición? El único método seguro es hacer una negación que la contradiga. Este método, sin embargo, nos devuelve al procedimiento de la Tabla 16.1. Y por tanto, a menudo hacemos una negación inadecuada. Intentamos encontrar un contraejemplo para una conclusión pero buscamos el ejemplo erróneo.

* * *

La utilización de contraejemplos en los silogismos se ha examinado de forma experimental. Recordará usted que un silogismo es un argumento sobre conjuntos de individuos, p. ej.:

Todos los artistas son apicultores.

Todos los químicos son apicultores.

Por tanto, todos los artistas son químicos.

Esta inferencia no es válida, y una forma de demostrar su no validez es pensar en un contraejemplo. Por ejemplo, hay dos tipos de apicultores, los que son artistas y los que son químicos. En este caso, todos los artistas son apicultores y todos los químicos son apicultores, pero en contra de la conclusión ningún artista es químico.

¿De qué otra forma podríamos establecer que una inferencia no es válida si no utilizamos un contraejemplo? Las teorías psicológicas actuales basadas en reglas formales no hacen uso de los contraejemplos, y la no validez se establece, o bien probando que las premisas implican la negación de la conclusión, o, en el caso en que una conclusión es al menos consistente con las premisas, al no poder encontrar una prueba formal de que la conclusión se sigue de las premisas. Un inconveniente de este procedimiento es que nos llevaría un largo tiempo buscar todas las posibles pruebas. Otro inconveniente, como han señalado los lógicos, es que si no conseguimos encontrar una prueba siempre penderá sobre nosotros una duda sobre si nuestra búsqueda fue completa. Por el contrario, un contraejemplo es un antimodelo “que salta a la vista” y que demuestra la no validez.

Algunas de las personas con un razonamiento pobre extraen casi siempre una conclusión de las premisas silogísticas que se les proponen, sean cuales sean. No han conseguido buscar modelos alternativos para ellas. Otras de esas personas

responden casi siempre que no se sigue nada de las premisas silogísticas que tienen modelos múltiples. No han sido capaces de captar que los modelos apoyan una conclusión común (véase la Tabla 10.3).

No solamente son las teorías basadas en reglas formales de inferencia las que ignoran el papel de los contraejemplos, sino que al menos hay una descripción basada en modelos mentales que también lo hace. Thad Polk y el difunto Allen Newell propusieron esta teoría, que utiliza modelos alternativos, pero no contraejemplos. Su teoría proporciona una buena explicación para las diferencias en el comportamiento silogístico entre un sujeto y otro. Pero sospecho que algún otro proceso de la teoría lleva a cabo el trabajo de los contraejemplos. Una razón para mi sospecha es un hallazgo de un estudio más reciente. Un experimento exteriorizó el proceso de búsqueda de contraejemplos. A los participantes se les dieron silogismos con conclusiones putativas, p. ej.:

Algunos de los chefs son músicos.

Ninguno de los músicos es pintor.

Por tanto, ninguno de los chefs es pintor.

Todos los silogismos se referían a chefs, músicos y pintores y a los participantes se les dieron unos sencillos recortables de papel para representarles –sombreros de chef para representar a los chefs, guitarras para representar a los músicos y paletas para representar a los pintores–. Unos sencillos “palotes” representaban a los individuos, y los participantes podían añadir sombreros, guitarras o paletas con el fin de representar las premisas. Su tarea era construir con los recortables un modelo externo con el fin de refutar la conclusión, o bien decir que la tarea era imposible si pensaban que la conclusión se seguía de las premisas. El experimento examinó veinte silogismos, cuatro con conclusiones válidas y dieciséis con conclusiones no válidas.

Los participantes fueron capaces de construir contraejemplos correctos para el cincuenta y nueve por ciento de las inferencias no válidas –un resultado que era mucho mejor que el comportamiento al azar, porque la mayoría de los modelos de las premisas no son contraejemplos–. Todos los participantes fueron capaces de construir contraejemplos correctos, pero el rango en su actuación variaba desde el noventa y cinco por ciento correcto en el caso del mejor de los participante hasta solo un veinticinco por ciento correcto para el peor. Estos márgenes de capacidad son característicos en el razonamiento silogístico de la población en general.

Como predice la teoría de modelos, una causa importante de error era que a menudo los participantes no captaban qué constituía un contraejemplo para ciertos tipos de proposición. Por ejemplo, con conclusiones de la forma: algunos de los chefs no son pintores, unas veces construían un modelo en el que algunos de los chefs eran pintores, y otras veces un modelo en el que ninguno de los chefs era pintor. El contraejemplo correcto es uno en el que todos los chefs son pintores. Se hizo una evaluación estricta de los participantes, y por eso estos errores no contaban como construcción de contraejemplos.

Un aspecto llamativo de los resultados era la gran variedad de procedimientos diferentes que desarrollaban los participantes. Los críticos han sostenido que un fallo grave de la teoría de modelos era su fracaso al especificar cómo buscamos los contraejemplos. El problema no era conceptual, dado que un programa de ordenador que implementaba la teoría contenía un procedimiento que creaba contraejemplos. La dificultad estaba, en cambio, en descubrir cómo llevamos a cabo la búsqueda nosotros. El experimento con los recortables mostró que podemos construir contraejemplos. En realidad, para construir nuevos tipos de individuos, utilizamos los mismos procedimientos que el programa, pero se produce una discrepancia grave entre programa y persona. El programa seguía un único procedimiento fijo para encontrar los contraejemplos; los participantes, no. Normalmente, la operación que utilizaban para buscar modelos alternativos era la de añadir nuevos individuos a los modelos. Pero resultaba imposible predecir si añadirían uno o dos individuos nuevos o el orden en que los añadirían.

Quienes no tienen formación en lógica pueden construir contraejemplos para refutar conclusiones no válidas que se les dan a evaluar. Pero, ¿piensan en contraejemplos al extraer sus propias conclusiones? Un segundo experimento comparaba el razonamiento con y sin el uso de recortables. No se produjo

diferencia en cuanto a exactitud entre las dos situaciones. Recordará usted que algunos silogismos dan lugar a una conclusión válida a partir de un único modelo, mientras que otros dan lugar a una conclusión válida solo a partir de modelos múltiples. Los que exigían solo un modelo eran más fáciles que los que exigían modelos múltiples. Para los que requieren un modelo, la mayor parte de las conclusiones de los participantes se basaban en un modelo externo único utilizando los recortables. No obstante, todos los sujetos construyeron modelos externos múltiples para algunos silogismos, variando desde dos participantes que los construyeron en las tres cuartas partes de los problemas, hasta uno que los construyó en menos de una décima parte. Un resultado crucial era que, para todos los sujetos, era más probable que lo hicieran en el caso de los silogismos que exigían modelos múltiples. Pero el experimento no mostró que estuvieran buscando contraejemplos. Simplemente, podía que hubieran estado explorando diferentes posibilidades: un problema con los silogismos es, por supuesto, que sus premisas están abiertas a muchas interpretaciones. Vuelvo ahora a algunos estudios sobre el razonamiento con conectivas proposicionales que corroboran el uso de contraejemplos.

* * *

La conclusión de una inferencia no válida no necesita ser cierta ni siquiera si las premisas son ciertas, y por tanto hay dos tipos de inferencia no válida. En uno, la conclusión es inconsistente con las premisas. Es decir, no se cumple en ninguna de las posibilidades compatibles con ellas, p. ej.:

O Dan está en Madrid o bien Bill está en Seúl, pero no ambas cosas.

Por tanto, Dan está en Madrid y Bill está en Seúl

Con estas oraciones sencillas sobre diferentes personas en diferentes ciudades, ni

el significado ni el conocimiento general modulan la interpretación lógica de la conectiva. Y por tanto, la premisa es consistente con dos posibilidades, pero ambas son inconsistentes con la conclusión, y en consecuencia no se puede seguir de las premisas.

Un segundo tipo de inferencia no válida es más sutil. La conclusión es consistente con las premisas, pero no se sigue de ellas. Es decir, se cumple al menos en una posibilidad consistente con las premisas, pero no en todas ellas, p. ej.:

Dan está en Madrid o Bill está en Seúl o ambas cosas.

Por tanto, Dan está en Madrid y Bill está en Seúl.

La premisa da lugar a tres posibilidades, y la conclusión se cumple en una de ellas –aquella en la que ambos sujetos están en sus respectivas ciudades–. Pero la conclusión no es una inferencia válida, porque dos posibilidades compatibles con la premisa son sus antimodelos: Dan está en Madrid y Bill no está en Seúl, y Dan no está en Madrid y Bill está en Seúl. Aquellos que al razonar pueden construir un modelo para una de estas posibilidades y reconocerlo como un contraejemplo saben que la inferencia no es válida: las premisas pueden ser ciertas, pero la conclusión falsa.

En resumen, una conclusión que es inconsistente con las premisas no se cumple en ninguno de los modelos de las premisas. Pero otro tipo de conclusión no válida se cumple al menos en un modelo de las premisas, pero no en todos ellos. Por tanto, debería ser más probable que utilizáramos contraejemplos para estas inferencias. La teoría predice también que debería ser más fácil inferir que una conclusión no es válida cuando es inconsistente con las premisas que cuando es consistente con ellas. Cualquier modelo de las premisas es un antimodelo de una conclusión inconsistente con ellas, pero ningún modelo de las premisas es un antimodelo de una conclusión consistente con ellas. Si, como sucede a menudo, no logramos considerar todos los modelos de las premisas, deberíamos seguir estando en lo cierto al evaluar una inferencia inconsistente, porque su conclusión

está en conflicto con cualquier modelo de las premisas. Pero corremos el riesgo de equivocarnos con una conclusión consistente, porque podemos pasar por alto un modelo que sea su contraejemplo. Se sigue que las conclusiones inconsistentes deberían dar lugar a un mayor número de evaluaciones correctas que las conclusiones consistentes pero no válidas.

Varios experimentos han corroborado estas predicciones, y sus resultados proporcionan la mejor prueba de que al razonar se utilizan contraejemplos. Los participantes evaluaban inferencias que se les daban, y para descubrir si captaban la fuerza de los contraejemplos, tenían que escribir las justificaciones de sus juicios. Dos jueces diferentes evaluaron las justificaciones escritas y estuvieron de acuerdo en casi todas ellas. Solo incluían una justificación en el grupo de las que utilizaban un contraejemplo si el participante describía un antimodelo –una posibilidad en la que las premisas eran verdaderas pero la conclusión era falsa–. Los estudios mostraban que los participantes utilizaban métodos variados, incluidos contraejemplos y contradicciones. Los resultados corroboraban las predicciones de la teoría de modelos. Los participantes utilizaban más a menudo los contraejemplos para refutar conclusiones consistentes con las premisas (cincuenta y uno por ciento de los problemas) que para refutar conclusiones inconsistentes con ellas (veintinueve por ciento de los problemas) –en este caso, tendían por el contrario a señalar la contradicción–. Los participantes fueron más exactos al reconocer las conclusiones no válidas cuando eran inconsistentes con las premisas que cuando eran consistentes con ellas. En general, el uso de contraejemplos correlacionaba con las evaluaciones correctas.

La tarea de escribir justificaciones para las evaluaciones podría haber afectado a los juicios de los participantes. Por tanto, en otro experimento se comparó a los participantes que escribían justificaciones con aquellos que no lo hacían. No hubo diferencia en la exactitud de sus evaluaciones. Aquellos que escribían justificaciones tendían, como antes, a utilizar contraejemplos para refutar conclusiones consistentes con las premisas, y a detectar la contradicción con las conclusiones inconsistentes con las premisas. La semejanza en la exactitud y la velocidad de respuesta sugiere que los participantes de los dos grupos estaban utilizando procedimientos similares. Eran capaces de utilizar contraejemplos para refutar conclusiones no válidas tanto si tenían que justificar sus evaluaciones como si no.

* * *

Más de la mitad de las personas de la habitación hablan francés.

Más de la mitad de las personas de la habitación hablan inglés.

¿Se sigue que más de la mitad de las personas de la habitación hablan francés e inglés?

Rechazamos la inferencia como no válida. Pero, ¿imaginamos un contraejemplo? Esta cuestión fue respondida por un experimento. Los participantes evaluaron una serie de inferencias partiendo de una, dos o tres premisas, siendo todas ellas afirmaciones referidas a “más de la mitad” de un conjunto dado. Las conclusiones se basaban o bien en “al menos una” o bien en “más de la mitad” de ellas. Un ejemplo de inferencia difícil con tres premisas es:

Más de la mitad de los pacientes tienen shnupfosis.

Más de la mitad de los pacientes tienen hustosis.

Más de la mitad de los pacientes tienen frognosis.

¿Se sigue que más de la mitad de los pacientes tienen al menos dos de las tres enfermedades?

Los participantes tenían papel y lápiz y podían escribir notas o dibujar diagramas, y también debían pensar en voz alta mientras trataban de resolver las inferencias. Como en los experimentos anteriores, desarrollaron diversos métodos, pero para más de las tres cuartas partes de los problemas dibujaron diagramas de las premisas. En los diagramas típicos representaban el conjunto al que se refería el tema de cada premisa con un cierto número de señales, como por ejemplo X. Luego representaban los otros conjuntos con señales que suponían solo un poco más de la mitad de las que tenía el conjunto del tema e intentaban minimizar la superposición entre los conjuntos. Resumiendo, buscaban un contraejemplo. La Figura 16.1 muestra un contraejemplo típico para la primera de las inferencias anteriores con la conclusión de que más de la mitad de las personas que están en la habitación hablan francés e inglés. Se basa en un diagrama que dibujó un participante. Este participante pretendía minimizar la superposición de los dos conjuntos, y el diagrama refuta la conclusión. Sin embargo, no refuta la conclusión:

Al menos alguna persona habla inglés y francés.

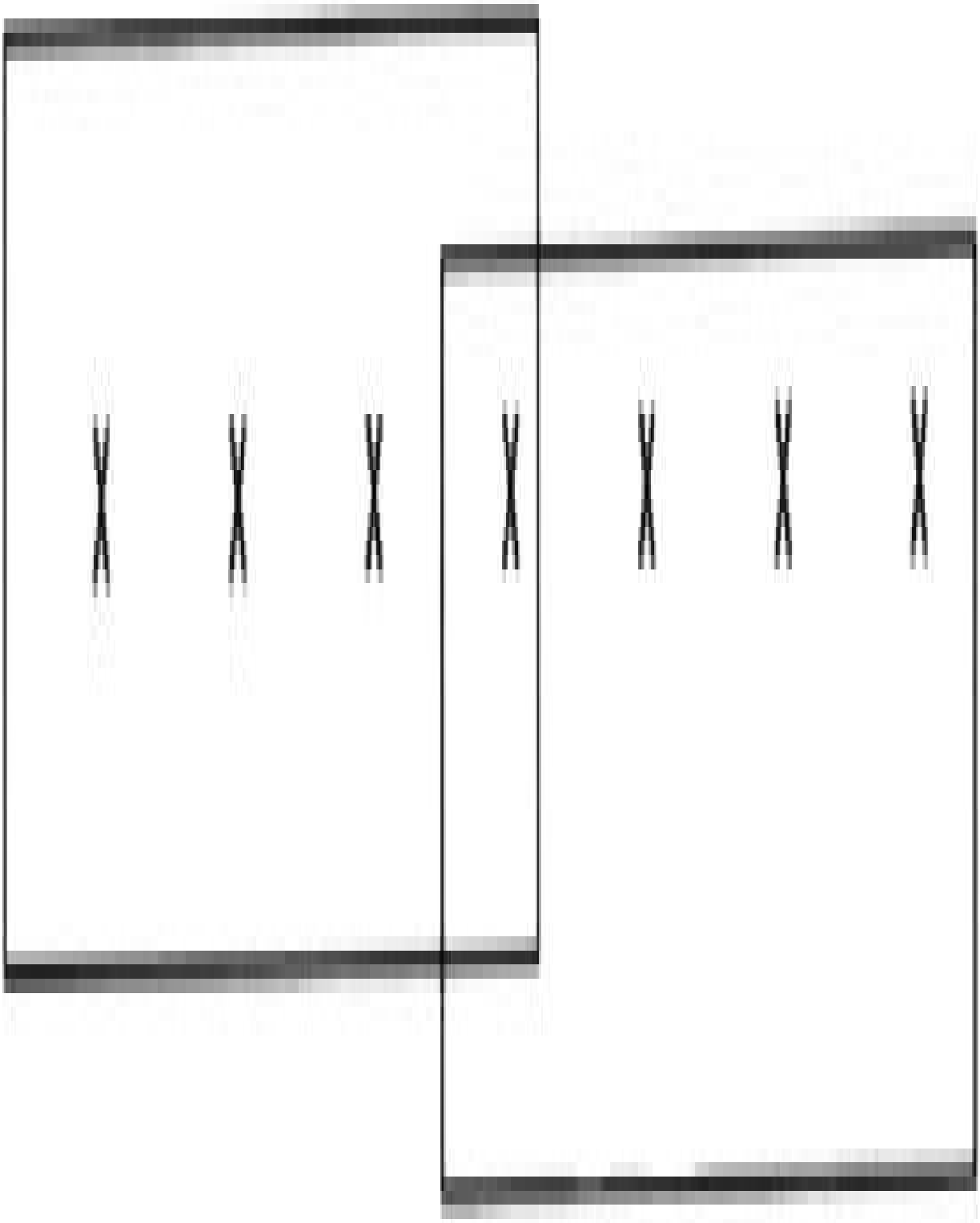


Fig. 16.1. Un diagrama de un contraejemplo basado en un dibujo hecho por un participante en un experimento. Cada “x” representa a una persona. Las premisas eran que más de la mitad de las personas de la habitación hablan francés, y que más de la mitad de ellas hablan inglés. El diagrama es un contraejemplo para la conclusión de que entre ellas, más de la mitad hablan ambos idiomas.

Ningún diagrama correcto lo hace, y por tanto esta conclusión es válida.

Los protocolos revelaban cómo trataban los participantes de resolver los problemas y mostraban que construían contraejemplos en el sesenta por ciento de los ensayos con conclusiones no válidas. Su error típico era no encontrar un contraejemplo donde, en realidad, existía uno. Estos errores eran con frecuencia el resultado de la adopción, por su parte, de un método de búsqueda erróneo. Una vez que adoptaban un procedimiento para los problemas tendían a no cambiarlo. Las tres inferencias de la Tabla 16.2 ilustran este punto. El diagrama para refutar la primera conclusión no requiere que los químicos se separen en dos grupos, mientras que esta separación es necesaria para un antimodelo de la segunda conclusión. Los posibles contraejemplos para la primera conclusión incluyen a los que refutan la segunda conclusión, y por ello la primera conclusión debería ser más fácil de refutar que la segunda. La tendencia a minimizar las superposiciones, que es el procedimiento más frecuente, no logra dar lugar a un contraejemplo para la tercera conclusión. Esa es la razón por la que este problema es tan difícil. Tenemos que maximizar la superposición de los tres conjuntos, dado que la superposición supone menos de la mitad del conjunto de personas (véase el diagrama 3 en la tabla). Por ello, la dificultad del problema viene de una aplicación persistente, pero errónea, de un método satisfactorio en otro caso.

Tabla 16.2. Tres inferencias no válidas basadas en “más de la mitad”

Las premisas:

Más de la mitad de las personas son artistas.

Más de la mitad de las personas son apicultores.

Más de la mitad de las personas son químicos.

1. ¿Se sigue que más de la mitad de las personas son artistas, apicultores y químicos? (los veinte participantes dijeron: no).

2. ¿Se sigue que al menos una persona es artista, apicultor y químico? (solo ocho de los participantes dijeron: no).

3. ¿Se sigue que más de la mitad de las personas tienen al menos dos de las tres ocupaciones: artista, apicultor, químico? (solo un participante dijo: no).

Hay tres tipos de diagramas que son antimodelos que refutan las inferencias anteriores:

1. Este diagrama refuta la primera conclusión: Más de la mitad de las personas son artistas, apicultores y químicos:

Artista Apicultor

Artista Apicultor

Artista Apicultor Químico

Químico

Químico

2. Este diagrama refuta las conclusiones primera y segunda: al menos una persona es artista, apicultor y químico:

Artista Químico

Artista Químico

Artista Apicultor

Apicultor

Apicultor Químico

3. Este diagrama refuta las conclusiones primera y tercera: más de la mitad de las personas tienen al menos dos de las tres ocupaciones: artista, apicultor y químico:

Artista Apicultor Químico

Artista Apicultor Químico

Artista

Apicultor

Químico

* * *

Una de las molestias para los experimentadores que estudian los contraejemplos es que tienen que utilizar un procedimiento especial a fin de detectarlos. Tienen que pedir a los participantes que piensen en voz alta, o que escriban justificaciones para sus evaluaciones, o que construyan modelos externos con recortables. Una posible forma de evitar esta molestia es encontrar una región del cerebro que se active al razonar, pero solo cuando se está tratando de

encontrar un contraejemplo. Un estudio sin publicar que utiliza la resonancia magnéticta funcional (RMf) –un procedimiento que ya describí con anterioridad (en el Capítulo 2)– afrontaba esa cuestión.

El procedimiento de RMf exige repetidas observaciones, y por tanto el experimento no podía utilizar inferencias basadas en, “más de la mitad”, porque las respuestas resultarían pronto conocidas. En lugar de ese tipo de problemas, utilizó otros con números y contenidos que variaban de un ensayo a otro, como por ejemplo:

Hay cinco estudiantes en una habitación.

Tres o más de esos estudiantes son corredores.

Tres o más de esos estudiantes son escritores.

Tres o más de esos estudiantes son bailarines.

¿Se sigue que al menos uno de los escritores que están en la habitación es estudiante?

Esta conclusión en concreto se sigue inmediatamente a partir de la tercera premisa, y por tanto no hay necesidad de buscar contraejemplos. Pero, consideremos esta conclusión alternativa:

¿Se sigue que al menos uno de los estudiantes que están en la habitación es las tres cosas: corredor, escritor y bailarín?

La conclusión es consistente con las premisas, pero no se sigue de ellas, y por tanto su evaluación correcta, suponiendo que se usen modelos, exige la búsqueda de contraejemplos. En un estudio previo al de RMf, los participantes tuvieron que evaluar las inferencias y llevar a cabo test mentales aritméticos basados en una presentación independiente de las mismas premisas. Un ejemplo típico de test aritmético basado en las premisas anteriores era:

X es el número de estudiantes. ¿Es (X + 24) + (X – 1) = 29?

Las premisas anteriores establecen que hay cinco estudiantes, es decir, X = 5, y por tanto un cálculo sencillo muestra que la respuesta a la pregunta es: No. El estudio preliminar hizo posible asociar la dificultad relativa de los dos tipos de inferencia –las fáciles que no exigían una búsqueda de contraejemplos y las difíciles que sí lo hacían– con la dificultad relativa de dos tipos de problema de cálculo mental. El ejemplo anterior era un problema fácil de cálculo.

El estudio RMf mostraba que una vez que los participantes habían entendido las premisas, las regiones del cerebro activas durante el razonamiento y durante los cálculos eran independientes una de otra. El razonamiento implicaba regiones del hemisferio cerebral derecho –en los lóbulos frontales, mientras que el cálculo mental implicaba regiones del hemisferio izquierdo–. La Figura 16.2 muestra el resultado más llamativo. Una región que los expertos conocen como “polo frontal derecho” estaba activa solo cuando un problema de razonamiento exigía la búsqueda de contraejemplos. Las otras fluctuaciones en los niveles de activación de la figura no dieron lugar a ningún resultado significativo. Un estudio no es decisivo, pero sí sugiere que la activación en el polo frontal derecho puede estar detrás de una búsqueda de contraejemplos. Si es así, los experimentadores pueden evaluar conjeturas sobre contraejemplos sin exigir a los participantes que cuenten en voz alta sus pensamientos, que escriban justificaciones o que dibujen diagramas.

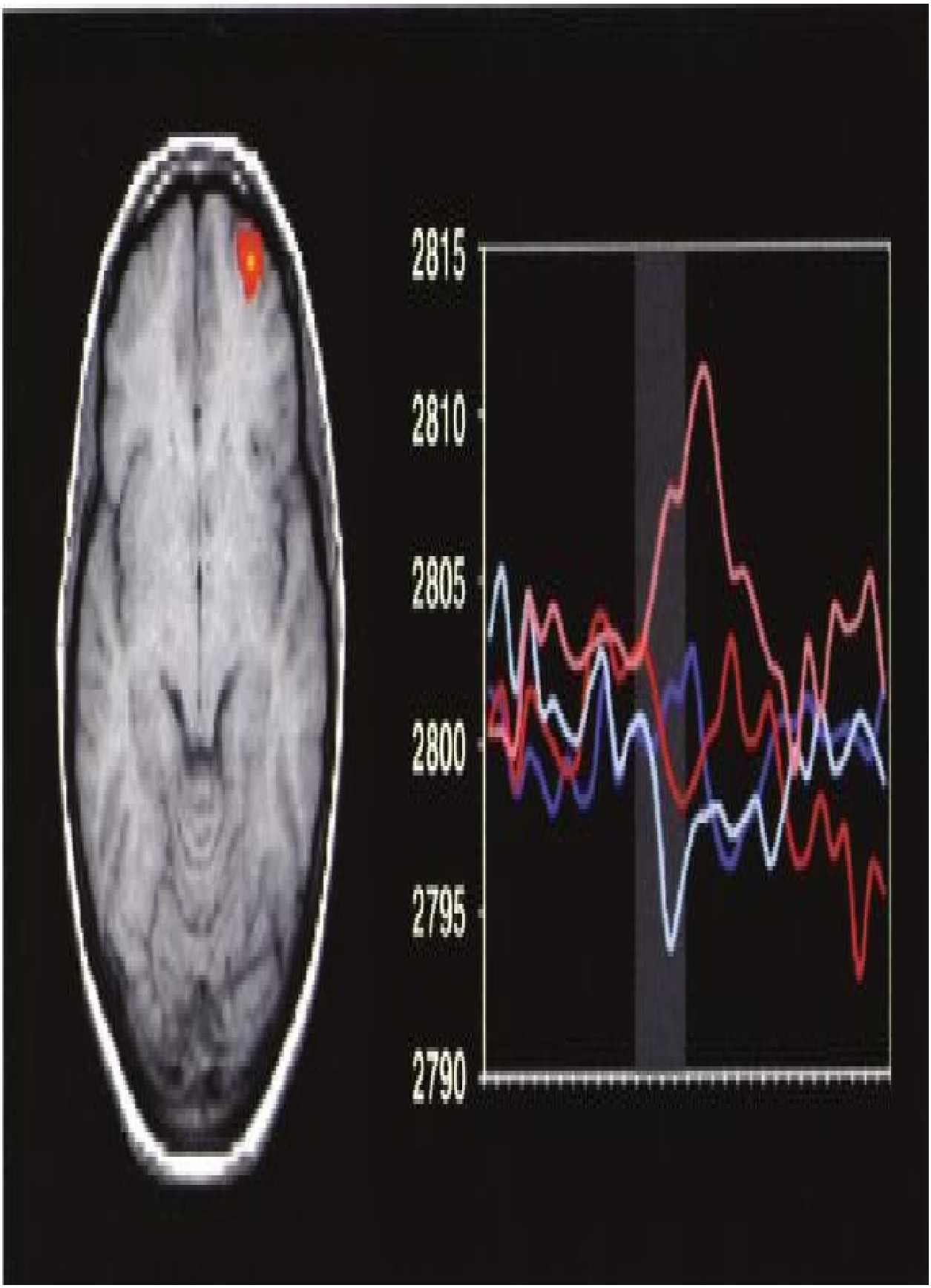

2815
2810
2805
2800
2795
2790

Fig. 16.2. A la izquierda aparece una imagen de un corte horizontal del cerebro en el que la parte frontal se sitúa en la parte superior de la página. La región en rojo es una zona del lóbulo frontal derecho que se conoce como el "polo frontal". El gráfico de la derecha muestra la cantidad de activación en esta región para cuatro situaciones diferentes en un experimento: el rosa muestra la activación para problemas de razonamiento en los que los participantes buscaron contraejemplos, el rojo muestra la activación para problemas de razonamiento en los que los participantes no buscaron contraejemplos, el azul claro muestra la activación para problemas de cálculo fáciles y el azul oscuro muestra la activación para problemas de cálculo difíciles. El eje horizontal es el tiempo, y la banda vertical gris corresponde a los ocho segundos durante los que se mostró el problema.

* * *

Captamos la fuerza de los contraejemplos, pero podemos tener dificultades para captar su importancia cuando formulamos nuestras propias conclusiones. Podemos tener dificultad para establecer el cotraejemplo correcto para una conclusión. Por ejemplo, podemos pensar que un contraejemplo para la conclusión, algunos de los artistas son químicos, es un antimodelo de una posibilidad en la que algunos de los artistas no son químicos. Ambas posibilidades podrían ser ciertas, sin embargo, y por tanto no son inconsistentes una con la otra. Sin embargo, incluso si no tenemos formación en lógica, podemos construir contraejemplos al razonar con conectivas proposicionales y al razonar con cuantificadores como por ejemplo: "más de la mitad" y "al menos tres".

La tendencia a utilizar contraejemplos varía de una persona a otra, y hay quien los utiliza raras veces. Son varios los factores que parecen dar lugar a esta variación. Un factor es la capacidad intelectual: una señal de inteligencia es el uso correcto de los contraejemplos. Correlaciona con las evaluaciones correctas de las conclusiones. Otro factor es el que una conclusión no válida sea o no consistente con las premisas. Tendemos a refutar las conclusiones inconsistentes

detectando la contradicción entre ellas y las premisas. Sin embargo, las conclusiones consistentes no se pueden refutar de esa forma, y por tanto tendemos a imaginar contraejemplos. También es probable que nuestra experiencia afecte a la utilización de contraejemplos. Podemos haber adquirido pericia en imaginarlos como resultado de enfrentarnos a problemas de un cierto tipo. Hay otro factor importante. Cuando sacamos nuestras propias conclusiones, tenemos que basarlas en los modelos de las premisas, y este procedimiento puede dificultar la búsqueda de antimodelos.

* * *

¿Cuál es el criterio apropiado para evaluar la racionalidad de nuestro razonamiento deductivo? La mayoría de las teorías presuponen que es la validez –que una conclusión debe ser verdadera en caso de que las premisas lo sean–. Pero hay dos psicólogos, Mike Oaksford y Nick Chater, que no están de acuerdo. Han sostenido que el criterio correcto para evaluar nuestro razonamiento cotidiano y nuestras deducciones en el laboratorio es el cálculo de probabilidades. Defienden incluso que este cálculo describe nuestras deducciones. Creen que nuestro razonamiento es una adaptación racional al mundo, y que depende de un proceso abarcable. Debido a que el cálculo de probabilidades es inabarcable, añaden una cláusula adicional; utilizamos heurísticos en vez de las intrincadas computaciones que exige el cálculo. Sin embargo, rechazan los resultados de Tversky y Kahneman de que los heurísticos nos pueden llevar a errores en el pensamiento probabilístico. Abogan en lugar de eso por heurísticos diferentes, que ellos afirman que son racionales. Su teoría es ingeniosa, y, aceptando ciertas suposiciones sobre las interpretaciones de las premisas, hace una buena descripción de las frecuencias con que los participantes extraen diversas conclusiones silogísticas. Por ejemplo, dadas las premisas:

Algunos de los autores son panaderos.

Algunos de los panaderos son jugadores de cricket.

hay algunas personas que extraen la conclusión:

Algunos de los autores son jugadores de cricket.

No es válida, pero se sigue del cálculo de probabilidades. Y por tanto cuando razonamos, según esta versión, extraemos conclusiones que son posibles y a veces además necesarias.

La mayoría de los participantes en los experimentos que he descrito anteriormente construían un contraejemplo para el silogismo anterior. Si el cálculo de probabilidades es la norma adecuada para evaluar el razonamiento, eran irracionales al hacerlo. Pero esta afirmación es extraña. El dilema para la teoría es que hay quien rechaza las deducciones no válidas y quien no. Si hay razón para rechazarlas, se es irracional al aceptarlas, y si hay razón para aceptarlas, entonces se es irracional al rechazarlas. Ninguna descripción consistente de la racionalidad puede tenerlo todo.

En realidad, el cálculo de probabilidades proporciona una pobre descripción de nuestro razonamiento deductivo. No somos conscientes de que pensamos en probabilidades cuando hacemos deducciones silogísticas o proposicionales. No las mencionamos cuando pensamos en voz alta. Estudios recientes han mostrado que cuando evaluamos si una conclusión es necesariamente verdadera, probablemente verdadera o posiblemente verdadera, nuestra actuación varía de una tarea a otra. He aquí un ejemplo sencillo para agudizar el contraste:

Dentro de la caja hay una canica roja o una canica verde, o ambas.

Por tanto, dentro de la caja hay una canica roja y una canica verde.

¿Cree usted que la conclusión se sigue de la premisa? Lo dudo. Pero se sigue en el cálculo de probabilidades, y por tanto en la descripción probabilística se debería aceptar que la conclusión se sigue de las premisas. En realidad, los participantes en los experimentos con conectivas proposicionales que describí anteriormente en este capítulo utilizaron espontáneamente contraejemplos para refutar inferencias de este tipo. Por tanto, la validez es central en su razonamiento. Los lógicos y los psicólogos, incluidos los mismos Oaksford y Chater, entienden también la noción de validez, y por eso podemos estar seguros de que al menos algunos la tenemos. En realidad, la habilidad para encontrar el criterio de validez deductiva correlaciona con medidas de inteligencia. Cuanto más inteligentes somos, más conclusiones válidas extraemos y más nos abstenemos de deducciones no válidas. El grado en que nuestras deducciones producen conclusiones probables en lugar de producir conclusiones válidas es una medida de nuestra estupidez.

* * *

Nuestro razonamiento cotidiano se inicia con frecuencia fuera de nuestra consciencia. Es rápido, automático, y a menudo nos lleva a conclusiones verdaderas. Como señaló Peirce, hacemos conjeturas correctas. Pero, ¿es racional? Algunos psicólogos dicen que debe serlo porque tendemos a conseguir nuestras metas, y distinguen entre un tipo de racionalidad para el razonamiento cotidiano y otro tipo para las deducciones deliberadas. Pero el mero hecho de alcanzar una meta no es prueba de ningún tipo de racionalidad a menos que uno esté dispuesto a asignar racionalidad a todas las criaturas vivientes. Las abejas, por ejemplo, han existido desde millones de años antes que nosotros. Alcanzan sus metas, y sin embargo sus preferencias por las fuentes de néctar violan un principio lógico sencillo. Los experimentos han mostrado que en ciertos casos prefieren la fuente A a la fuente B, y la fuente B a la fuente C, y sin embargo prefieren la fuente C a la fuente A. De manera similar, nuestra habilidad explicativa nos lleva a flagrantes inconsistencias con el cálculo de probabilidades. ¿Por qué, por ejemplo, mujeres muy inteligentes se casan con hombres menos inteligentes? Como usted recordará, en lugar de reconocer un

caso de regresión a la media ofrecemos explicaciones. Y algunas veces encontramos más probables las conjunciones que las proposiciones que las constituyen. Mejor reconocerlo y decir que también nosotros cometemos errores.

El capítulo empezó con tres preguntas que ahora puedo responder. Primero, sí reconocemos la fuerza de los contraejemplos. Captamos el principio de que una inferencia válida no tiene contraejemplos. Esta comprensión, sin embargo, no nos vacuna contra los errores. A menudo sobrevivimos –aunque no siempre– a pesar de nuestros errores irracionales. Lo hacemos todavía mejor cuando somos racionales al razonar. Segundo, la mejor receta para provocar contraejemplos parece ser pedirnos que evaluemos la conclusión de otra persona, utilizar premisas sencillas que provoquen modelos múltiples de posibilidades, y asegurarse de que la conclusión se cumple al menos en uno de estos modelos. Finalmente, la teoría de modelos parece explicar nuestra actuación. Es probable que adoptemos una actitud crítica frente a la conclusión de otro, y cuando no tenemos que formular la conclusión por nosotros mismos, esta tiende a no llenar nuestra memoria de trabajo. Es más probable que susciten contraejemplos las conclusiones no válidas que son consistentes con las premisas, mientras que aquellas que son inconsistentes con las premisas es más probable que provoquen la detección de la contradicción. No tenemos un procedimiento fijo para buscar contraejemplos, y este capítulo ha mostrado que cuando los buscamos cometemos errores predecibles.

CAPÍTULO 17

Verdades, mentiras y razonamiento superior

¿Qué es la verdad? Poncio Pilatos planteó la pregunta a Jesús (San Juan 18:37), pero Jesús no contestó. El burlón Pilatos no esperaba una respuesta, escribió el filósofo del siglo dieciséis Francis Bacon. La pregunta nos llega con tanta mística que muchos suponen, como Pilatos, que es demasiado profunda para ser respondida. Sin embargo, sus perfiles generales no son tan misteriosos. Si preguntamos a la gente, "¿qué significa decir que una proposición es verdadera?", la respuesta más frecuente es: la proposición describe una situación que se corresponde con los hechos. Esta teoría de la "correspondencia", que tiene una larga historia filosófica, es correcta. Una proposición es verdadera cuando corresponde a cómo es el mundo, y es falsa cuando no corresponde a cómo es el mundo. Todos los sistemas legales se colapsarían si no aceptáramos este principio.

La dificultad para formular la descripción de la correspondencia está en poner en palabras qué quiere decir que la correspondencia se cumpla. Los lógicos separan el lenguaje para hablar sobre lógica, al que se refieren como "metalenguaje", de la propia lógica, que tiene su propio lenguaje, el lenguaje "objeto". Por tanto, en lógica, la solución a la correspondencia es decir, por ejemplo, que la oración "la nieve es blanca" es verdadera si la nieve es blanca. Lo que parece insustancial, pero que realmente ilustra la diferencia entre una afirmación en el metalenguaje: "la nieve es blanca" es verdadero, y una afirmación en el lenguaje que se está analizando, el lenguaje objeto: la nieve es blanca. Los lógicos hablan de las oraciones como verdaderas o falsas, pero recordará usted que en un capítulo anterior distinguí entre una oración como por ejemplo "yo puedo oírte ahora", y las muchas proposiciones que expresa dependiendo de quién lo dice, a quién se lo dice, y el contexto en que se pronuncia. Por tanto, en un lado de la correspondencia hay una proposición expresada mediante la utilización de una oración en un contexto determinado, y en el otro lado de la correspondencia hay una situación del mundo. Puedo intentar identificar una proposición determinada de esta manera: la proposición expresada por la oración "yo puedo oírte ahora"

cuando Pat O’Brian se la dijo a Viv Waugh por teléfono a las 9 de la mañana del lunes 10 de enero de 2005 en la esquina de Bleecker Street y Broadway en Manhattan. La proposición es verdadera si realmente Pat podía oír a Viv en esa ocasión. Esta afirmación parece ser correcta, pero es difícil generalizarla. Quizá lo más que uno puede decir es que una proposición se refiere a ciertos individuos o entidades, y les adscribe propiedades o relaciones. Es verdadera cuando esos individuos o entidades tienen las propiedades o relaciones adscritas. Imagino que cuando los robots puedan percibir y describir el mundo, este principio estará plasmado en su software.

El poeta romántico John Keats escribió en una carta a un amigo, “todavía no he sido nunca capaz de percibir cómo se puede saber que algo es verdadero mediante razonamiento consecutivo –y sin embargo debe saberse–”. Pero, admitiendo que un conjunto de proposiciones sea verdadero, cualquier otra proposición que sea su consecuencia válida también es verdadera. Y este hecho resuelve la perplejidad de Keats respecto a cómo el “razonamiento consecutivo” lleva a la verdad. Lo hace siendo sensato, es decir, siendo una inferencia válida a partir de premisas verdaderas. Algunos filósofos han elevado la visión inferencial a una teoría de la verdad hecha y derecha, la llamada teoría de la “coherencia”, según la cual una proposición es verdadera si es coherente con otras proposiciones que se sabe que son verdaderas. Pero la coherencia parece ser parasitaria del razonamiento consecutivo a partir de verdades de correspondencia.

La comunicación funciona debido a que una convención fundamental es que quien habla dice la verdad. Sin esta convención, el discurso no tendría sentido. Quienes escuchan tendrían que comprobar la verdad de cada proposición. De manera similar, sin esa convención las mentiras no funcionarían. Si embargo, solo de vez en cuando podemos comprobar la veracidad de la comunicación. Como resultado, podemos descubrir que una proposición es verdadera. Entonces sabemos que es verdadera, y podemos a su vez testificarlo. No obstante, tomamos como verdaderas la mayoría de las proposiciones.

La verdad nos importa porque las comunicaciones sobre el mundo tienen más valor para nosotros, y son más útiles para dirigir nuestras acciones si son verdaderas que si son falsas. Por supuesto, si sabemos que una proposición es falsa, entonces su negación es verdadera, y las negaciones pueden dar información. A menudo, sin embargo, no sabemos si las proposiciones son verdaderas. Pero, si son verdaderas, sabemos cómo es el mundo, y lo sabemos

por poderes. Es decir, no hemos sido testigos del estado del mundo, pero alguien lo ha sido, y lo ha descrito para nosotros. Cuando dos jóvenes ingenieros de Chernobyl fueron a investigar la suerte del reactor después de la explosión, vieron que había resultado destruido, y volvieron a la sala de control a describir lo que había pasado. Los dos jóvenes pagaron más tarde sus observaciones muriendo por enfermedades debidas a la radiación, pero su descripción convenció finalmente a los ingenieros a cargo de la central de que el reactor había resultado destruido. Los ingenieros lo notificaron a las autoridades de Moscú y finalmente se inició la evacuación de la ciudad, unas horas después del accidente. Una descripción verdadera salvó vidas.

* * *

La verdad es más sencilla en lógica que en la vida real. En un cálculo lógico, una oración es verdadera cuando se corresponde con algún dominio de interpretación; en caso contrario, es falsa. Si conocemos los valores de verdad de un conjunto de proposiciones elementales, en principio es simple determinar el valor de verdad de una oración compuesta construida con esas proposiciones mediante conectivas lógicas. Los lógicos han generalizado esta idea para mostrar cómo determinar los valores de verdad de las oraciones en un lenguaje lógico que contiene cuantificadores, es decir, los análogos lógicos a “cualquier” y “algún”.

Un asunto delicado respecto a las proposiciones en la vida real es si toda proposición tiene que ser verdadera o falsa. El significado de la negación funciona de manera sencilla si se da por hecha esta asunción. Como señaló Aristóteles, si una proposición afirmativa es verdadera, entonces su negación es falsa, y si una proposición afirmativa es falsa, entonces su negación es verdadera. Por ejemplo, si es verdadero que usted está leyendo este libro, entonces es falso que usted no está leyendo este libro. Pero ciertas oraciones dan lugar a proposiciones problemáticas. Yo le digo a usted: Mi mujer ha dejado de fumar. Pero, supongamos que mi mujer no ha fumado nunca en su vida. ¿Mi proposición es verdadera o falsa? Si usted dice que es falsa, entonces su negación sería verdadera: mi mujer no ha dejado de fumar. Pero esta proposición implica que sigue fumando. Otra proposición problemática, que Bertrand Russell

hizo famosa, es: “El actual rey de Francia es calvo.” No hay actual rey de Francia. Por tanto, ¿la proposición es verdadera o falsa? Si usted dice que es falsa, entonces de nuevo su negación debe ser cierta: “El actual rey de Francia no es calvo.” Los lógicos están preparados a veces para aceptar estas desagradables consecuencias. En la vida cotidiana, sin embargo, algunas oraciones dependen de presuposiciones para expresar proposiciones. La proposición sobre mi mujer fumando presupone que sí fumó en algún momento. La oración sobre el rey de Francia presupone que hay un rey de Francia. Si estas presuposiciones son falsas, entonces ni las oraciones ni sus negaciones expresan proposiciones. No son ni verdaderas ni falsas. Esta afirmación no implica la existencia de un tercer tipo de valor de verdad, más allá de la verdad y la falsedad. Las oraciones no consiguen tener un valor de verdad, lo mismo que una división por cero no da lugar a ningún resultado.

* * *

Si nuestra esposa nos dice, mirando por la ventana, “Hay un unicornio en el jardín”, podemos decidir si la proposición es o no verdadera mirando también por la ventana. Entendemos la proposición: nuestro sistema interpretativo construye un modelo mental de la posibilidad que describe. Un procedimiento de verificación compara nuestro modelo de la proposición con el modelo perceptivo de la escena que construye nuestro sistema visual. Los mecanismos que subyacen en este proceso no se entienden bien, pero lo que tienen que hacer es hacer corresponder [map] los referentes de nuestro modelo de una posibilidad con objetos de nuestro modelo perceptivo, y determinar si las propiedades y las relaciones de un modelo se corresponden con las del otro. Los modelos de las proposiciones expresadas mediante el lenguaje son rudimentarios comparados con los modelos perceptivos del mundo, que contienen mucha más información –muchos más referentes, propiedades y relaciones–. Por tanto, si la proposición es verdadera, el modelo basado en la proposición se corresponde con parte del modelo basado en la percepción: todos los referentes, propiedades y relaciones del modelo proposicional se corresponden con los del modelo perceptivo que contiene todavía muchas más cosas.

El proceso es falible. Podemos malinterpretar la proposición y construir el

modelo equivocado; podemos percibir erróneamente el mundo y construir el modelo equivocado; nuestro mapping puede ser erróneo e identificar mal un referente, una propiedad o una relación. Pero el proceso es correcto con la suficiente frecuencia como para que nos sintamos seguros respecto a si una proposición es verdadera o si es falsa. Muchas proposiciones, lamentablemente, no se pueden percibir ni como verdaderas ni como falsas. Un ejemplo es la proposición: El presidente George W. Bush tiene un rancho. No podemos percibir su valor de verdad. Como máximo podemos percibir pruebas pertinentes para su valor de verdad, p. ej., muestra el registro de su compra. Por tanto, no podemos saber que tiene un rancho, pero podemos llegar a creer la proposición.

* * *

“Verdadero” y “falso” son términos que pueden referirse a la relación entre las proposiciones y el mundo. Por tanto, una proposición como:

Es falso que David Thomson sea un crítico de teatro

afirma una proposición sobre una proposición. Es una proposición metalingüística. Ciertas inferencias se refieren a afirmaciones sobre la verdad o falsedad de las proposiciones. Consideremos, por ejemplo, este breve diálogo:

Galileo dice: me retracto y aseguro que la Tierra es el centro del universo.

El inquisidor dice: Eso es mentira.

Inferimos que si el inquisidor es sincero entonces cree que Galileo no se ha retractado o no asegura que la Tierra es el centro del universo o ninguna de las dos cosas.

Las inferencias a partir del discurso sobre la verdad y la falsedad son comunes en la vida real. Pero pueden dar lugar a profundas paradojas. La más famosa es la paradoja del “mentiroso” a la que alude San Pablo en su Primera Epístola a Tito:

Uno de ellos, su propio profeta, dijo, los cretenses son siempre mentirosos, bestias malvadas, panzas holgazanas. Este testimonio es verdadero.

Pablo parece estarse refiriendo a Epiménides, un profeta cretense, y estar haciendo un juicio moral más que una afirmación sobre las paradojas del lenguaje. Sin embargo, la implicación es que, como cretense, Epiménides está mintiendo cuando dice que los cretenses son siempre mentirosos. Y por tanto su afirmación es falsa. No se sigue una paradoja real, sino solamente que los cretenses no mienten siempre. Una genuina versión paradójica de la paradoja del “mentiroso” la condensa en una sola oración referida a ella misma:

Esta oración es falsa.

Si la oración anterior es verdadera, entonces es falsa; si es falsa, entonces es verdadera. Los lógicos detestan las oraciones de este tipo, lo mismo que detestan cualquier paradoja que amenaza la consistencia de un sistema. Algunos analistas sitúan la raíz de la paradoja del “mentiroso” en el hecho de que la oración se refiere a ella misma. Pero la autorreferencia no es siempre paradójica (véase la primera –y única– nota a pie de página de este libro). Las paradojas aparecen en el lenguaje natural porque es su propio metalenguaje. Es decir, podemos utilizar oraciones en un idioma para hablar de las relaciones entre las oraciones en ese idioma y el mundo. Como señalé anteriormente, los lógicos mantienen el lenguaje para hablar sobre una lógica, el metalenguaje, separado del lenguaje de

la propia lógica. De esta forma, evitan la paradoja del mentiroso en lógica.

Debido a que un lenguaje natural es su propio metalenguaje, podemos razonar sobre las consecuencias de la verdad y la falsedad. Por ejemplo, dado el par de oraciones:

Es falso que la siguiente oración sea verdadera.

Adolf vive en Argentina.

la primera oración implica que la segunda oración es falsa, y por tanto se sigue que Adolf no vive en Argentina. Como usted puede imaginar, estos problemas pueden ser muy complicados. Se parecen a las notas al margen de los libros de texto del colegio (el Shorter Latin Primer en mi época), como por ejemplo: "Pasar a la página 183." Cuando pasamos a esa página, nos da otra instrucción para pasar a otra página, y así una y otra vez –hasta que llegamos a una página con un mensaje grosero, o nos envía de nuevo a la página 183, de forma que podemos entrar en un bucle infinito pasando de una página a otra.

Los problemas metalingüísticos se refieren a menudo a individuos que afirman la verdad o falsedad de sus propias afirmaciones, p. ej.:

Lance dice: yo estoy mintiendo y Art está mintiendo.

Art dice: Lance está mintiendo.

¿Quién está diciendo la verdad?

Lance no puede asegurar que está mintiendo sin que sea una paradoja a menos que asegure algo más en conjunción con su afirmación. Pero sí que afirma algo más –que Art está mintiendo–. Suponiendo que esta segunda cláusula sea falsa, la conjunción de Lance como un todo es falsa, y por tanto su afirmación de que está mintiendo es verdadera. Se sigue que Art, la segunda persona que habla, no está mintiendo sino diciendo la verdad, y por tanto está diciendo la verdad porque afirma que Lance está mintiendo. El razonamiento es un poco mareante pero, si lo piensa detenidamente, verá usted que es cierto.

Este tipo de problemas se formulan a menudo como problemas de “caballeros y villanos”. He aquí el mismo problema con este disfraz:

Hay dos tipos de personas: caballeros y villanos. Los caballeros siempre dicen la verdad, mientras que los villanos siempre mienten.

Lancelot dice: yo soy un villano y Arthur es un villano.

Arthur dice: Lancelot es un villano.

¿Cuál es el estatus de Lancelot y Arthur: caballero, villano, o imposible de decir?

La solución es análoga al problema anterior: Lancelot es un villano y Arthur es un caballero.

El primer psicólogo que estudió estos problemas fue Lance Rips. Postulaba que utilizamos un único procedimiento para resolverlos, y los planteaba como un reto a la teoría de modelos. De acuerdo con su versión, suponemos que, en un problema, el primer individuo está diciendo la verdad, y luego utilizamos procedimientos formales para extraer tantas inferencias como sea posible.

También podemos suponer que el primer individuo está diciendo una mentira y perseguir las consecuencias de esta asunción. Cada etapa inferencial se basa en un procedimiento formal. Para enfrentarse con estos problemas, sin embargo, Rips adopta algunas reglas específicas del contenido para hacer inferencias sobre caballeros (y qué es verdadero) y sobre villanos (y qué es falso). El sistema se describe en la Tabla 17.1. Se apoya en los resultados según los cuales quienes razonan cometen más errores y tardan más en resolver los problemas que, en sus derivaciones, requieren más etapas formales. Sin embargo, resulta irónico aplicar procedimientos formales a asuntos de verdad o falsedad. Los métodos formales para la lógica se oponen a los métodos basados en la verdad y falsedad, y como mencioné en el Capítulo 1, existe entre ellos una brecha insalvable.

Tabla 17.1. Las reglas formales de Rips para los problemas de caballeros y villanos

Los procedimientos formales para razonar se complementan con reglas para enfrentarse a los problemas de caballeros y villanos. Consideremos la inferencia válida:

A dice que B es un villano. A es un caballero. ∴ B es un villano.

Esta regla formal da lugar a la inferencia:

1. x dice: p

x es un caballero

∴ p

He simplificado aquí la sintaxis de Rips. Asimismo, para derivar la inferencia válida:

A dice que B es un villano. A es un villano. ∴ B no es un villano.

la teoría tiene esta regla:

2. x dice: p

x es un villano

∴ no p

También hay reglas para inferir que si x no es un caballero entonces x es un villano, y que si x no es un villano entonces x es un caballero.

He aquí un ejemplo de una prueba formal que utiliza el procedimiento único de Rips basado en suposiciones. El problema es:

A dice: B es un villano.

B dice: A es un caballero si y solo si C es un caballero.

¿Qué es C?

La respuesta se deriva de esta prueba:

Supongamos que A es un caballero.

∴ B es un villano [a partir de la regla anterior]

∴ No (A es un caballero si y solo si C es un caballero) [a partir de la regla 2 anterior]

∴ C no es un caballero [a partir de la suposición y la línea anterior]

∴ C es un villano [a partir de la regla que las interdefine]

En este punto, el programa hace una nueva suposición:

Supongamos que A no es un caballero.

∴ A es un villano.

∴ B no es un villano.

∴ B es un caballero.

∴ A es un caballero si y solo si C es un caballero.

∴ C no es un caballero.

∴ C es un villano.

A partir de ambas asunciones respecto a A, se sigue que C es un villano, y por tanto C es un villano (y el programa de Rips tiene una regla general para hacer una inferencia de este tipo).

La teoría de modelos propone que no utilizamos un único método para resolver estos problemas. Ciertamente, la mayoría de las personas llegan al laboratorio sin un método previo en relación con ellos y tienen que aprender sobre la marcha. Los primeros esfuerzos son dubitativos. Pero pueden desarrollar una diversidad de procedimientos para enfrentarse con los problemas, o al menos con aquellos que están dentro de su competencia. Los problemas concretos que encuentran influyen por tanto en los métodos concretos que desarrollan. Uno de los métodos, como el que se propone en la explicación de Rips, utiliza suposiciones.

Un método diferente depende de premisas que son “circulares”, es decir, aquellas que refutan el estatus de aquellos que las hacen, como en el ejemplo anterior:

Lancelot dice: Yo soy un villano y Arthur es un villano.

Las premisas circulares atrapan la atención. Casi todo el mundo capta que parecen autorrefutarse. Muchas personas no pueden ir más lejos, pero algunas sí se dan cuenta de que la premisa de Lancelot no puede ser verdadera como un todo. Debe ser un villano. Suponiendo, entonces, que su premisa es falsa, su primera cláusula es verdadera, y por tanto su segunda cláusula debe ser falsa. Por tanto, se sigue que:

Lancelot es un villano y Arthur es un caballero.

Un tercer método puede desarrollarse a partir de encuentros con premisas como:

Gawaine asegura que Percivale es un villano.

Bedivere asegura que Percivale es un villano.

Percivale asegura que Gawaine es un caballero y Bedivere es un villano.

Podemos darnos cuenta de que como Gawaine y Bedivere aseguran la misma proposición, o ambos son villanos o ambos son caballeros. Percivale, sin embargo, les sitúa en categorías diferentes. Por tanto, Percivale debe ser un villano. Se sigue que tanto la proposición de Gawaine como la de Bedivere son verdaderas: son caballeros. El método detecta también la situación en que dos individuos afirman proposiciones opuestas sobre el mismo individuo y determina la falsedad de cualquier premisa que considere que los dos individuos

pertenecen al mismo grupo.

Ruth Byrne y sus colegas han mostrado que los participantes sí que desarrollan, en el curso de un experimento, sus propios métodos para resolver estos problemas. Se vuelven más rápidos y más exactos incluso sin recibir ninguna información sobre la exactitud de su actuación. Aquellos que desarrollan un método que es aplicable a un margen reducido de problemas lo hacen peor cuando encuentran problemas a los que el método no es aplicable. Los que hacen uso del método general basado en suposiciones tienen menos dificultades cuando encuentran problemas hechos a medida para otro método. No tienen que aprender un nuevo método para resolverlos, sino que pueden continuar utilizando el método más general. Y diferentes participantes desarrollan métodos diversos. Pueden seguir las consecuencias de las suposiciones, pueden entender las consecuencias de las premisas circulares o pueden darse cuenta de ciertos patrones interesantes de las premisas.

Los problemas de caballeros y villanos están sacados de libros de problemas. No se producen en el mundo cotidiano, y por eso es fácil desecharlos. Pero debería usted recordar que son solo una forma abreviada de plantear argumentos, como por ejemplo:

Yo mentía cuando te dije que Pat estaba mintiendo.

Ella dijo que había enviado la carta y que yo era un mentiroso.

El uso de modelos para el razonamiento proposicional no puede hacer frente en solitario a proposiciones de este tipo. Para resolver la verdad y la falsedad exigen una habilidad de nivel superior. Esta habilidad es esencial para el desarrollo de la lógica formal como disciplina intelectual, debido a que nos hace capaces de relacionar patrones formales con asuntos de verdad y falsedad.

Los psicólogos pueden ir más allá de los problemas metalingüísticos para formular problemas que son metalógicos, p. ej.:

Hay dos tipos de personas: los lógicos que siempre hacen deducciones válidas y los políticos que nunca hacen deducciones válidas. Abe dice que o bien Bob está diciendo la verdad o bien Bob es un político, pero no ambas cosas. Bob dice que Abe está diciendo la verdad. Cal infiere que Bob es un político. ¿Es Cal un lógico o un político?

Para resolver este problema, vamos a empezar con la suposición de que Abe está diciendo la verdad. Se sigue que hay dos posibilidades: Bob está diciendo la verdad y no es un político, o Bob no está diciendo la verdad pero es un político. En la primera posibilidad, de lo que Bob dice se sigue que Abe está diciendo la verdad. Lo cual es consistente con la suposición y apoya la conclusión de que Bob no es un político. En la segunda posibilidad, de lo que dice Bob se sigue que Abe no está diciendo la verdad, lo que es inconsistente con la suposición de que sí lo está haciendo. Por tanto, esta segunda posibilidad se puede eliminar, y sabemos que a partir de la suposición de que Abe está diciendo la verdad se sigue que Bob no es un político. Ahora, vamos a hacer la suposición de que Abe no está diciendo la verdad. Una posibilidad es que Bob esté diciendo la verdad y sea un político, pero él asegura que Abe está diciendo la verdad. Lo cual es inconsistente con la suposición, y por tanto podemos eliminar esta posibilidad. La otra posibilidad es que Bob no esté diciendo la verdad ni sea político. Ahora, a partir de la falsedad de la proposición de Bob se sigue que Abe no está diciendo la verdad. Lo cual es consistente con la suposición. Y por tanto, sabemos que a partir de la suposición de que Abe no está diciendo la verdad se sigue que Bob no es un político. Ambas suposiciones, que son exhaustivas, dan lugar a la conclusión válida de que Bob no es un político. Sin embargo, Cal infiere que Bob es un político, y por tanto la inferencia de Cal no es válida. En consecuencia, Cal es un político. Sencillo, ¿no?

* * *

Con frecuencia razonamos sobre nuestro propio razonamiento o el razonamiento de los demás. Una circunstancia en la que llevamos a cabo este

metarrazonamiento, o “razonamiento superior”, como yo lo llamo, es cuando nos preocupa lo que pensarán los demás. Si estamos dando vueltas por un lugar público esperando encontrarnos con alguien entonces podemos preocuparnos de que los demás sospechen que no estamos haciendo nada bueno. Por tanto, para disipar sus sospechas, miraremos nuestro reloj y fruncíremos el ceño de forma llamativa de manera que los otros infieran que estamos esperando a alguien que llega tarde a una cita, y que nuestros paseos no tienen intenciones maliciosas. Ervin Goffman consideró estos casos como conducta pública diseñada para suscitar ciertas inferencias en los demás. Los magos, como mencioné en un capítulo anterior, están siempre tratando de imaginar cómo razonará su público, y tratando pensar en las posibilidades que no inferirá. Todo el tema de la cooperación con otras personas, o la competición con ellas, gira también en torno a la realización de inferencias sobre lo que ellos inferirán.

George Erdos ha examinado en un experimento el metarrazonamiento, utilizando un famoso problema:

Tres mujeres, Ann, Beth y Cath, se sientan en fila una delante de otra. La sombrerera tiene cinco sombreros: tres sombreros blancos y dos sombreros negros. Pone un sombrero sobre la cabeza de cada mujer sin que ella vea el color de su propio sombrero. Pero Ann puede ver los sombreros de Beth y Cath, y Beth puede ver el sombrero de Cath. Por supuesto, Cath no puede ver el sombrero de nadie. Las mujeres saben lo que ha hecho la sombrerera. Son honestas, y pueden razonar de forma válida. La sombrerera les pide que digan, “sí” si conocen el color del sombrero que llevan puesto, y “no” en caso contrario. Ann dice, “No”. Beth dice, “No”. ¿Qué dice Cath y por qué?

Para explicar cómo resolvemos el problema y cómo llevamos a cabo el metarrazonamiento en general, la teoría de modelos necesita un componente adicional. Este componente capacita a nuestro sistema de razonamiento para construir un modelo respecto a cómo razonan otros individuos. Consideremos, por un momento, cómo resolvemos nosotros el problema.

El primer paso es entender la situación inicial. Hay cinco sombreros, tres blancos y dos negros. Tres personas llevan, cada una, uno de los sombreros. Por

consiguiente, hay siete posibilidades diferentes: todas las mujeres llevan un sombrero blanco, Ann y Beth llevan sombreros blancos y Cath lleva un sombrero negro, etc. Para resolver este problema tenemos que razonar sobre cómo razonó cada una de las mujeres. Ann razonó: si yo veo dos sombreros negros, entonces llevaría un sombrero blanco, porque en el conjunto de cinco solo hay dos sombreros negros. Pero, puesto que dijo, “no”, no sabía el color de su sombrero, y por tanto no veía dos sombreros negros. Tanto Beth como Cath hacen esta inferencia. Por tanto, ambas infieren que Ann ve al menos un sombrero blanco, y por tanto hay tres posibilidades. O Beth lleva un sombrero blanco o Cath lleva un sombrero blanco, o las dos lo llevan. Beth sigue razonando a partir de estas tres posibilidades: si veo que Cath lleva un sombrero negro, entonces yo debo llevar un sombrero blanco. Pero, puesto que dijo: “no”, no veía un sombrero negro. Cath hizo esta inferencia, y sigue razonando: llevo un sombrero blanco, y por tanto respondo, “sí”.

Este problema es difícil. No es fácil razonar sobre el razonamiento de los demás, porque tenemos que construir modelos de sus modelos. Y este problema en concreto nos exige pensar sobre modelos de tres posibilidades distintas cuando razonamos sobre el razonamiento de Beth y Cath. Tres posibilidades son difíciles de mantener presentes.

Los problemas de sombreros y análogos se pueden generalizar para cualquier número de individuos. Por ejemplo, existe un problema de sombreros sobre cuatro individuos y cuatro sombreros blancos y tres sombreros negros, y un problema de sombreros sobre 100 individuos y 100 sombreros blancos y 99 sombreros negros, etc. Existe también un problema sobre dos individuos y tres sombreros: dos blancos y uno negro. En este caso, Ann dice, “No”, y Beth puede inferir que Ann no ve el sombrero negro, porque si lo hubiera hecho, habría inferido que ella llevaba un sombrero blanco. Ergo Beth lleva un sombrero blanco. Este problema es más fácil, como mostró Erdos en un experimento, porque solo necesitamos considerar cómo razonó un individuo respecto a cómo razonaba otro, y las consecuencias dan lugar a un único modelo para una posibilidad –equivalente a la etapa final del problema de tres personas.

Para burlar a un oponente, o conseguir lo mejor de la cooperación con un compañero, tendemos a pensar sobre lo que es probable que ese otro individuo haga. Este tipo de pensamiento se conoce como estratégico, y aparece permanentemente en los negocios, la política y la vida diaria. Si usted lee la transcripción de las cintas de Kennedy, que registran las conversaciones entre el

presidente Kennedy y sus consejeros durante la crisis de los misiles cubanos en octubre de 1962, se sorprenderá por la frecuencia con que los participantes intentaron comprender lo que estaba pasando por la mente del primer ministro soviético Khrushchev y por la de sus consejeros. Un comentario típico al inicio de la crisis describe cómo razonaba el entonces jefe de la CIA, John A. McCone:

El razonamiento de Mr. McCone... era: Si esto es así, entonces, ¿qué posible razón tienen ellos [los soviéticos] para entrar en Cuba de la forma en que lo han hecho, con misiles tierra- aire y misiles tipo crucero? Simplemente no podía entender por qué los soviéticos estaban reforzando tan intensamente la postura defensiva de Cuba. Debía haber algo detrás. Lo que le llevaba entonces a creer que debían estar trayendo MBRM (misiles balísticos de alcance medio).

El problema era por qué Krushchev había ordenado enviar MBRM a Cuba. ¿Era para persuadir al presidente Kennedy de aflojar sus lazos con Berlín, para ampliar la cobertura de los misiles soviéticos, para desanimar a los Estados Unidos respecto a una invasión de Cuba –la razón más visible de Krushchev, o una consecuencia de discusiones entre distintas facciones dentro del Kremlin–? Ahora sabemos que fue una decisión de Krushchev, pero eso solo hace más difícil saber cuál de las restantes posibilidades le motivó. Como señalan Ernest May y Philip Zelikow, los editores de las transcripciones de las cintas, “Kennedy y sus consejeros entendieron el razonamiento del Kremlin mejor de lo que lo hicieron muchos expertos que escribieron sobre la crisis a posteriori”.

La teoría que describe cómo debemos razonar estratégicamente es la teoría matemática de juegos, que incluye el teorema de John Nash sobre las estrategias optimas para todos los jugadores, basadas en predicciones racionales de las elecciones de otros jugadores. Pero, como señaló cuatro años más tarde en una reunión internacional Reinhard Selten, coganador del premio Nobel de Economía con Nash en 1994, “la teoría de los juegos es teología racional”. Lo que quiere decir es que en la vida diaria a menudo somos irracionales, y por tanto la teoría de juegos resulta pobre para predecir cómo nos implicaremos en las interacciones estratégicas con los demás. El pensar sobre lo que los demás están pensando es verdaderamente difícil, porque exige una representación, no solo de nuestro estado mental, sino también del estado mental de la otra persona.

No existe una teoría psicológica general sobre cómo llevamos a cabo ese tipo de razonamiento, pero un resultado sorprendente se refiere a juegos experimentales sencillos. Usted tiene la oportunidad, pongamos por caso, de elegir la opción A o la opción B, y lo mismo ocurre con el otro jugador. Cada uno de ustedes gana una cierta cantidad de dinero dependiendo de su opción conjunta, pero uno puede ganar más que el otro. Si puede usted predecir la elección de su oponente, entonces usted puede hacer una elección más prudente. Sin embargo, un estudio mostraba que los participantes jugarían a estos juegos sin preguntar siquiera cuáles serían las ganancias del otro jugador para las diversas opciones.

* * *

Una de las habilidades no reconocidas a los participantes en experimentos psicológicos es que pueden entender las instrucciones del experimentador respecto a lo que se supone que tienen que hacer y (normalmente) llegar a entender cómo hacerlo. John Duncan, en algunos estudios sin publicar, ha mostrado que la capacidad para seguir instrucciones, cuando una tarea está incluida en otra, correlaciona con una medida de la inteligencia de los participantes. La tarea de nivel bajo era leer en voz alta exclusivamente las letras de una columna vertical de letras y números. Había dos columnas de este tipo y la tarea de nivel alto era cambiar de una columna a la otra en la dirección señalada por una flecha, si bien las flechas aparecían solo de vez en cuando. Los participantes inteligentes podían afrontar la tarea. Pero los participantes menos inteligentes sabían lo que se supone que tenían que hacer, pero olvidaban cambiar de una columna a la otra, una vez que estaban en acción.

La capacidad para entender instrucciones es llamativa en el caso de experimentos sobre razonamiento. La experimentadora nos dice, por ejemplo:

Su tarea en este experimento es proporcionar una premisa perdida para un argumento. Junto a una premisa dada, debe implicar la conclusión.

Nos muestra un ejemplo:

Todos los críticos que admiran a Hitchcock piensan que su película Vértigo es estupenda.

Por tanto, David Thomson piensa que su película Vértigo es estupenda.

Nuestra tarea es proporcionar la premisa perdida que da lugar a la conclusión. Una vez que entendemos la tarea, no es demasiado difícil. La premisa perdida es:

David Thomson es un crítico que admira a Hitchcock.

La capacidad para entender las instrucciones en experimentos de razonamiento depende de que se entiendan ciertos conceptos, como premisa, conclusión, implicar, junto con nociones más prosaicas, incluidas las convenciones para un experimento. Estos conceptos y una capacidad para realizar razonamiento superior capacitan a los participantes para entender qué se espera que hagan. En el caso anterior, hay un conjunto de premisas, una conclusión, y una premisa perdida que hace válida una inferencia. Para un lógico, hay una solución trivial: hacer que la premisa perdida sea la propia conclusión. Las personas sensatas no dan esta respuesta en un experimento. Infieren que deberían pensar en una premisa que diga algo nuevo y dé lugar a la conclusión solo en conjunción con la premisa dada. También tienen que idear cómo realizar esta tarea. Tienen que utilizar su conocimiento para idear un método que les capacite en principio para resolver estos problemas. Tienen que razonar sobre el razonamiento. La premisa en el ejemplo anterior era: todos los críticos que admiran a Hitchcock piensan que su película Vértigo es estupenda. Por tanto, pueden imaginar un modelo de esta premisa en la que representan a todos los críticos que admiran a Hitchcock pensando que Vértigo es estupenda. Para asegurar que una persona como David

Thomson admira Vértigo es necesario representarle como un crítico que admira a Hitchcock. Por tanto, infieren que la premisa perdida es, por consiguiente, que David Thomson es un crítico que admira a Hitchcock.

Un reto interesante es diseñar un programa de ordenador que pueda inferir cómo llevar a cabo cualquier tipo de problema de razonamiento, por ejemplo, determinar si todas las oraciones de un conjunto podrían ser ciertas, proporcionar premisas que den lugar a una inferencia válida de forma parsimoniosa, evaluar si una conclusión dada se sigue de un conjunto de premisas, evaluar si una conclusión dada podría ser verdadera dado un conjunto de premisas, etc. Podemos entender cómo hacer cualquiera de estas tareas, y por tanto podemos razonar sobre el razonamiento. De manera similar, cuando llevamos a cabo una secuencia de inferencias deductivas, con frecuencia empezamos por darnos cuenta de algunas de sus propiedades formales. Por ejemplo, dada una secuencia de silogismos, podemos darnos cuenta de que cuando ambas premisas contienen el cuantificador “algunos” no se sigue nada. Este conocimiento formal incipiente surge del razonamiento sobre el razonamiento. Los principios formales que inferimos están muy por debajo de las inferencias que realmente podemos hacer, y algunas veces son erróneos. Por ejemplo, alguien afirmó después de un experimento sobre razonamiento condicional que “de dos premisas condicionales no se sigue nada”. La afirmación era verdadera en el experimento, pero es falsa en general. Los dos condicinales: Si Pat va a la fiesta entonces Viv no irá y si Pat no va a la fiesta entonces Viv no irá dan lugar a la inferencia válida de que Viv no irá. A pesar de todo, la observación de que podemos desarrollar nuestras propias reglas formales muestra que tenemos la capacidad de abstraerlas del contenido. Muestra también que no llegamos al experimento armados ya con esas reglas.

* * *

Todos entendemos qué significa decir que una proposición es verdadera: se corresponde con los hechos. Puesto que los lenguajes naturales incluyen términos como “verdadero” y “falso”, las proposiciones paradójicas se pueden enmarcar en ellos. Pero también podemos razonar a partir de premisas que incluyen esos términos metalingüísticos. La teoría de modelos se puede ampliar

para explicar estas inferencias, y, como en el razonamiento más sencillo, predice que podemos desarrollar varios métodos para enfrentarnos a ellas –una predicción que ha recibido corroboración experimental–. La memoria de trabajo restringe los métodos que podemos desarrollar para el razonamiento metalingüístico. Por ejemplo, los métodos que siguen las consecuencias de las asunciones solo lo hacen de manera limitada. Las tareas más complejas exigen razonamiento sobre el razonamiento. También razonamos de esa forma con el fin de entender las instrucciones para experimentos sobre razonamiento, y para llegar a entender cómo llevarlos a cabo.

La psicología tiene una propiedad que es única en ciencia. Necesita explicar su propio origen. Por tanto, una teoría sobre el razonamiento superior debería proporcionar algún insight sobre su propio desarrollo. La teoría de modelos postula una capacidad para pensar sobre la verdad o falsedad de las premisas, y sus imágenes indirectas bajo el disfraz de caballeros y villanos, que a su vez depende de una capacidad para reflexionar sobre problemas y procesos de pensamiento. Los científicos cognitivos puede utilizar esta capacidad para construir teorías del razonamiento. De manera similar los lógicos utilizan el mismo componente del pensamiento con el fin de crear cálculos formales, y reflexionar sobre las relaciones entre estos cálculos y sus significados.

Parte VI

Cómo desarrollamos la capacidad de razonar

... en una cultura que asigna un gran valor en la
imaginación y uno reducido a la lógica, bien podrían los
niños parecer más racionales que los adultos, porque un
niño no es, por naturaleza, nada más.

Auden

CAPÍTULO 18

Sobre el desarrollo

¿Cómo adquirimos la capacidad de razonar? La cuestión es tan desconcertante que nadie tiene una respuesta satisfactoria. La doctrina platónica es que todo es innato, incluidos todos nuestros conceptos y capacidades intelectuales, y que lo único que sucede en nuestro desarrollo es que las experiencias desencadenan ese conocimiento innato. Sin embargo, sí tenemos que aprender algunas cosas. Tenemos que aprender a leer, a montar en bicicleta, a cocinar. Tenemos que aprender que la Tierra es redonda, que gira y se mueve en torno al Sol, y que su clima se está haciendo más cálido. Se puede defender –y lo haré dentro de un momento– que en el desarrollo del lenguaje subyacen factores innatos. Pero las variaciones en la capacidad de razonamiento son mucho mayores que las variaciones en la capacidad lingüística. No me refiero en este caso a la elocuencia, sino a la mera competencia para producir y entender expresiones en nuestra lengua materna. Por tanto, la doctrina de las ideas innatas debe atemperarse un poco. No obstante, algunos filósofos modernos creen que nuestros conceptos, desde las bicicletas al pollo a la marengo son innatos, y solo esperan a que las experiencias adecuadas de la vida los desencadenen. El problema es que estos filósofos prueban demasiado. Sostienen que no podemos aprender conceptos, y por tanto estos deben ser innatos. Los conceptos innatos, sin embargo, deben haber evolucionado, y la evolución en sí misma es una forma de aprendizaje. Puede ser imitada en programas de ordenador que aprenden soluciones a problemas.

La antítesis es la pizarra en blanco de los empiristas. Según esta explicación nada es innato si se exceptúa un poderoso programa de aprendizaje, y nuestras experiencias ponen en marcha este programa, de forma que aprendemos todos nuestros conceptos y capacidades intelectuales. Ciertamente, el programa de aprendizaje puede capacitarnos incluso para aprender otros programas de aprendizaje más especializados. Aprendemos a aprender. El problema es que nadie ha explicado cómo funciona ese potente programa de aprendizaje. Un programa que simula la evolución sería demasiado ineficaz para explicar,

pongamos por caso, la rapidez con la que los niños aprenden el significado de las palabras. Un programa “conexionista” que simula el aprendizaje en una red de células cerebrales aprende asociaciones complejas. Pero estos programas hasta ahora no pueden aprender cómo se componen los significados de las oraciones a partir de los significados de las palabras, o cómo hacer algo que vaya más allá de las inferencias más simples –del tipo de las que defendí anteriormente como inconscientes.

* * *

Con frecuencia los padres suponen orgullosos que sus retoños aprenden a hablar. Los psicólogos, y los conductistas en particular, les animaban en esta opinión. La revolución de Chomsky respecto a la lingüística inspiró un giro de ciento ochenta grados. Según él sostenía, las oraciones no llevan su estructura gramatical escrita en la frente, los niños no reciben casi enseñanzas de gramática, y sin embargo adquieren su lengua materna a partir de la mera exposición al lenguaje. Deben existir por tanto restricciones innatas en el conjunto de gramáticas posibles para los lenguajes naturales. Estas restricciones constituyen una “gramática universal”, que proporciona una serie de opciones para cada uno de los componentes gramaticales de los lenguajes. Un componente, por ejemplo, se refiere al orden de sujeto, verbo y predicado; otro componente se refiere a si los adjetivos van delante o detrás de los nombres a los que modifican; etc. Por consiguiente, en términos de gramática, solo son posibles un número finito de lenguajes naturales. Los niños expuestos a su lengua materna solo tienen que identificar su opciones gramaticales concretas.

Las restricciones innatas también nos pueden capacitar para adquirir los principios que gobiernan el significado y la verdad. Los principios componen los significados de las oraciones a partir de los significados de sus componentes de acuerdo con las relaciones gramaticales existentes entre ellos. Nos capacitan para construir representaciones de a qué se refiere una oración, para imaginar lo que es posible de acuerdo con su significación, y para evaluar la verdad y falsedad a partir de una comparación de esos modelos con los modelos del mundo.

Podría usted preguntarse por qué se desarrolló el lenguaje. Una idea verosímil es que la capacidad para comunicarse dio a los seres humanos una ventaja en la lucha para sobrevivir y reproducirse. Algunos psicólogos han hecho esa sugerencia. Chomsky, no. Ciertamente, él sugiere que la función del lenguaje podría ser permitir, no que nos comuniquemos, sino que exterioricemos nuestros pensamientos. Puede que nunca sepamos qué condujo al desarrollo del lenguaje. Mientras tanto, la pizarra no puede estar completamente en blanco.

* * *

La síntesis correcta que explique el desarrollo intelectual debe combinar genes y experiencia. Jean Piaget, uno de los grandes pioneros de la investigación con niños, fue el primero que propuso este tipo de teoría del razonamiento. Su opinión era que una serie de procesos mentales, que culminan en torno a los doce años, da lugar a reglas de razonamiento que operan en forma de proposiciones. Un concepto central en su pensamiento era la noción de operación reversible. Puedo ilustrar esta noción con algunos experimentos piagetianos típicos sobre el tema de la “conservación”. El experimentador coge una bola de plastilina y le da la forma de un largo cilindro. ¿Sigue habiendo la misma cantidad de plastilina? Los niños de cinco o seis años dicen, “No, ahora hay más plastilina, porque es más larga”. El experimentador coge una jarra de agua y la vierte en un vaso más alto y más estrecho. Los niños dicen, “ahora hay más agua, porque llega más arriba”. El experimentador recoloca una hilera de seis vasos de forma que ahora están distribuidos en una hilera más larga. Los niños dicen, “ahora hay más vasos, porque la fila es más larga”. A diferencia de los adultos, los niños a esa edad no entienden que estas cantidades siguen siendo las mismas después de ciertos cambios. Tienen una imperfecta comprensión de la conservación.

La observaciones fueron sorprendentes. Nadie conocía la ignorancia de los niños sobre la conservación hasta que Piaget la demostró. Su trabajo impactó a los educadores, aunque él mismo no estaba interesado en la pedagogía. Decía que los niños cometían fallos en las pruebas de conservación porque todavía tenían que dominar el concepto de que ciertas operaciones se pueden invertir para volver al status quo. El experimentador podía volver a dar forma de bola a la

plastilina, verter de nuevo el agua en la jarra y volver a colocar los vasos en la hilera original.

Piaget hacía las preguntas correctas pero, viéndolo a posteriori, sus respuestas estaban equivocadas. Y su especulación era tendenciosa: buscaba una equivalencia entre las estructuras mentales y las estructuras matemáticas postuladas por un grupo de matemáticos que publicaba bajo el pseudónimo colectivo de “Bourbaki”. Nuestro desarrollo intelectual termina con esas estructuras, sostenía Piaget, y ellas explican cómo nos es posible dominar las matemáticas abstractas. Su teoría se resume en la Tabla 18.1.

Tabla 18.1. Teoría de Piaget sobre el desarrollo intelectual

• El conocimiento evoluciona gobernado por principios evolucionistas (¡lamarckianos!). Lo mismo los protozoos que la mente asimilan la realidad o se transforman a sí mismos para acomodarse a ella.

• El desarrollo es el resultado de una tendencia innata y automática para la autorregulación llamada “equilibración”. Cada nuevo equilibrio es el resultado de operaciones reversibles, pero es también una ocasión para la posterior corrección en caso de conflictos. El desarrollo según esto pasa por cuatro estadios principales:

1. Durante sus dos primeros años, los niños desarrollan una inteligencia “sensorio- motora” a partir de su acción sobre el mundo. Alcanzan el concepto de causación simple y el concepto de que los objetos que están fuera de la vista siguen existiendo.

2. Alrededor de los dos años, empiezan a adquirir el lenguaje. Interiorizan sus

acciones sobre el mundo, pero les falta el concepto completo de operaciones reversibles, y por tanto fallan en las tareas de conservación que describo en el texto. Solo tienen inteligencia “preoperacional”. No logran distinguir entre sus concepciones “egocéntricas” y la realidad.

3. Aproximadamente a los siete años, dominan las operaciones reversibles y el concepto de número, y por tanto alcanzan el estadio de las “operaciones concretas”. Conservan las cantidades.

4. Después de los once años, alcanzan el estadio de las “operaciones formales” en el que son capaces de elaborar hipótesis, calcular combinaciones y razonar utilizando una lógica formal que se corresponde con los significados lógicos de la negación, si, o e y (ver el Capítulo 8). Han adquirido tres tipos de grupos de estructuras mentales: grupos, relaciones seriales y espacios topológicos. Estas estructuras son fundamentales para las matemáticas (según Bourbaki).

▪

No siempre es fácil entender la teoría de Piaget, hace algunas predicciones erróneas, y algunos importantes defensores de las reglas formales la han abandonado. Al contrario de lo que dice la teoría, los bebés de menos de ocho meses comprenden que un objeto físico sigue existiendo cuando está fuera de la vista, los niños pueden hacer deducciones mucho antes de su undécimo cumpleaños, y, como mostré en un capítulo anterior, ni siquiera los adultos tienen un dominio completo de la lógica del “si”, “o” e “y”.

* * *

Si Piaget también está equivocado, ¿entonces qué pasa? El primer paso es definir qué se desarrolla. Las pruebas respecto al desarrollo de un conjunto de reglas formales para el razonamiento son cuanto menos equívocas. Pero, debido a que

eran la única oferta disponible, los psicólogos siguieron creyendo en ellas durante mucho tiempo. Algunos aún lo hacen. Por supuesto, todos podemos aprender reglas formales –en un curso de lógica, por ejemplo–. Pero la teoría de modelos postula, para nuestro desarrollo intelectual, un final de recorrido diferente:

El razonamiento es la capacidad de construir modelos a partir de la percepción, la descripción y el conocimiento, de formular conclusiones nuevas pero parsimoniosas a partir de esos modelos, y de captar la fuerza de los contraejemplos para esas conclusiones.

Voy a considerar esta descripción del razonamiento como la respuesta a qué es lo que se desarrolla. Y voy a esbozar una teoría de cómo se desarrolla. La teoría tiene tres componentes principales: las habilidades innatas de los niños, su adquisición de conocimiento tanto sobre el lenguaje como sobre el mundo y el aumento en la potencia computacional de su pensamiento.

* * *

Qué tienen los ricos que no tenemos nosotros, preguntó Scott Fitzgerald. Y Ernest Hemingway replicó: más dinero. Qué tienen los adultos que no tienen los niños, pregunta usted. Y yo replico: más conocimiento. Eso es verdad, aunque no es toda la verdad. La evolución ha equipado a los bebés con lo que necesitan para adquirir conocimiento. Pero también desarrollan la potencia computacional, y su crecimiento les capacita para hacer computaciones más complicadas. Imaginará usted que la forma de aumentar esta potencia es desarrollar algunos procedimientos nuevos y más sofisticados. En realidad, para cualquier computación basta con un reducido conjunto de procedimientos básicos y de formas de combinarlos. El aumento en la potencia computacional depende del desarrollo de una memoria de trabajo con una mayor capacidad. Por consiguiente, sospecho que las amplias diferencias en la capacidad de razonar entre una persona y otra son atribuibles a dos factores: diferencias de

conocimiento y diferencias en la potencia computacional.

* * *

La mayor parte de lo que hace falta para adquirir la habilidad de razonar puede estar “inscrito en el sistema nervioso”, es decir, ser consecuencia de los genes humanos. Los bebés tienen muchas más habilidades cognitivas que una pizarra en blanco. Los astutos investigadores en psicología han llegado a esta conclusión a partir del tiempo que los bebés pasan examinando diversos sucesos. Los sucesos no habituales atraen su atención durante más tiempo. Los experimentos tienen que asegurar que, después de que se ha producido un suceso, el estado final de las cosas les resulta familiar –no debería ser este estado lo que les sorprende por sí solo, sino la secuencia de sucesos que llevan hasta él–. Si un objeto se mueve de un lado a otro, se esconde durante un momento, y luego mágicamente aparece al otro lado de una barrera fija, los bebés de tres meses examinan la demostración durante más tiempo que si la barrera detiene el movimiento del objeto. Un objeto sólido no debería pasar a través de otro, y por tanto el primer suceso no es habitual. Pero, para que los bebés registren que el suceso no es habitual, deben compararlo con lo que es habitual. De hecho, esperan una posibilidad –que la barrera bloqueará al objeto–, pero esta posibilidad no se produce. Por tanto, los bebés perciben que es el mismo objeto el que se está moviendo antes y después de haber estado fuera de la vista durante un momento. Comprenden la permanencia de los objetos. Este resultado refuta la afirmación de Piaget: infería que los bebés no tienen el concepto de permanencia de los objetos, porque no buscan los objetos escondidos.

Durante el primer año, los bebés revelan que tiene en su mente todo un mundo de “física ingenua” –un mundo de sentido común en el que los objetos sólidos no pasan a través de agujeros más pequeños que ellos, no permanecen flotando en el aire si no se les sujeta y no se desvían de su trayectoria si nada choca con ellos–. Cuando el experimentador prepara trucos que violan estos principios, los bebés examinan durante más tiempo la demostración, lo mismo que los espectadores miran durante más tiempo al escenario después de que el mago David Devant hace desaparecer a la “Mascot Moth” (véase el Capítulo 3).

El conocimiento en física ingenua se desarrolla tan poco tiempo después del nacimiento que debe existir una predisposición innata para construirlo. Las experiencias son fundamentales, pero parecen ser los desencadenantes que evocan esta predisposición. Los bebés no necesitan aprender por ensayo y error que los objetos grandes no pueden pasar a través de agujeros pequeños. Asimismo, su maquinaria para comparar lo que sucede con lo que debería suceder debe ser innata. Tienen una capacidad innata para imaginar posibilidades.

¿Tienen los bebés la capacidad de razonar? El psicólogo Justin Halberda ha sostenido que a los dos años pueden hacer inferencias disyuntivas. Por ejemplo, ven una carrera entre dos animales y oyen una declaración positiva, “el ganador es el león”, o una declaración negativa, “el ganador no es el elefante”. En cualquiera de los casos, son capaces de extraer el ganador. En realidad, hacen la inferencia:

O el ganador es el león o el ganador es el elefante.

El ganador no es el elefante.

Por tanto, el ganador es el león.

En otra tarea, juegan a un juego en el que aprenden qué les gusta hacer a distintos animales, p. ej., al león le gusta nadar, al elefante le gustan las galletas, etc. En pruebas clave, ven un animal cuya preferencia conocen, como por ejemplo el león, y un animal que no habían visto antes, como por ejemplo un perro. Escuchan la oración, “me gusta la compota de manzana”, y se les pregunta, “¿quién dijo eso?”. No se les había dicho que a cada animal solo le podía gustar una cosa, pero todos los niños llegan a esa conclusión, e infieren que es al nuevo animal, el perro, al que le gusta la compota de manzana. Los niños están rechazando una alternativa con el fin de inferir la otra.

¿Razonan los bebés antes de entender el lenguaje? El lógico estoico Crisipo describía a un perro persiguiendo a su presa, al llegar a un punto en el que se unían tres caminos, olfateaba un camino, olfateaba otro camino, y luego inmediatamente echaba a correr por el tercer camino sin pararse a olfatearlo. El perro, afirmaba él, hizo una inferencia disyuntiva: no hay olor de la presa en los dos primeros caminos, y por tanto debe haberse ido por el tercero. Lamentablemente, según un reciente estudio los perros no tienen esa capacidad. Los bebés podrían ser capaces de hacer la inferencia disyuntiva, pero eso todavía tiene que demostrarse con un experimento. Alrededor de los doce meses, antes de haber adquirido el lenguaje, pueden hacer inducciones sencillas. Si ven a un adulto dando agua a un perro de juguete, imitarán su acción con otro animal, como por ejemplo un gato de juguete, pero no con entidades de otras categorías, como por ejemplo un coche de juguete. Tienen un concepto de animal y ese concepto restringe sus inferencias.

* * *

Los niños le atribuyen mente a otros seres humanos, a los animales e incluso a los objetos inanimados de sus juegos. Esta tendencia puede ser la base de los modelos que nos formamos de otros individuos, y de nuestras creencias sobre ellos, incluidas nuestras creencias sobre sus creencias. Pero nuestra capacidad para captar creencias sobre creencias tarda algún tiempo en desarrollarse. Este es un diálogo que ilustra esa tardanza:

Experimentador, mostrando al niño una caja de cerillas: ¿Qué crees que hay aquí dentro?

Niño: Cerillas.

El experimentador abre la caja para mostrar al niño que contienen M&M (conocidos también en el Reino Unido como “Smarties”). Llega una segunda niña, Pat.

Experimentador: ¿Qué pensará Pat que hay aquí dentro?

Niño: M&M.

Esta respuesta es típica del pensamiento egocéntrico de los niños de tres años. No pueden tener presente la diferencia entre lo que saben ellos y lo que saben otros. Pero un año más tarde, pueden dar la respuesta correcta. Separan su propia creencia de la falsa creencia del niño que no vio que la caja de cerillas contenía M&M.

Sin embargo, los niños de cuatro años siguen equivocándose respecto a las creencias de Pat sobre las creencias de otras personas –una creencia llamada de "segundo orden"–. Supongamos que usted es un niño de cuatro años que participa en el experimento. Comprende que Pat cree que la caja de cerillas contiene cerillas porque no vio al experimentador revelarle a usted su verdadero contenido. Si entra en la habitación una tercera niña, Viv, entonces Pat creerá que Viv también cree que la caja de cerillas contiene cerillas. Pero, supongamos que sin saberlo Pat usted ve al experimentador mostrarle a Viv que la caja de cerillas contiene M&M. Esta es la situación del juego:

Usted cree que la caja de cerillas contiene M&M.

Usted cree que Pat cree que contiene cerillas.

Usted cree que Viv cree que contiene M&M.

Usted cree que Pat cree que Viv cree que contiene cerillas.

Usted cree que esta creencia de segundo orden de Pat es falsa. ¿Por qué? Porque usted sabe que Viv cree que contiene M&M.

Esta es mucha información para tenerla presente. Por tanto, los niños no dominan las creencias de segundo orden, y su posible falsedad, hasta que tienen siete años. Hasta entonces, se inclinan a pensar que Pat cree, como ellos, que Viv cree que la caja de cerillas contiene M&M.

Entender las creencias sobre lo que otros creen parece depender de la capacidad para adscribir estados mentales a otras personas. Lo que a su vez parece tener una base innata, porque los niños autistas tienen afectada esta capacidad. Fallan en las pruebas sobre creencias de primer y de segundo orden. Sin embargo, pueden realizar inferencias del mismo tipo sobre mapas o dibujos. Pueden captar que una fotografía es antigua. Ya no se corresponde con la realidad lo mismo que no se corresponde con ella una creencia obsoleta. Para enfrentarse a las creencias de segundo orden, sin embargo, se necesita potencia computacional –una memoria de trabajo con una capacidad lo bastante grande como para mantener toda la información relevante–. Y una falta de capacidad puede ser la responsable de los problemas para entender las creencias sobre creencias. Ciertamente, algunos científicos sociales sostienen que el problema refleja la "maldición del conocimiento", es decir, nuestra dificultad para volver a capturar la inocencia intelectual una vez que hemos adquirido conocimiento sobre una situación.

* * *

La forma de aumentar el poder computacional es aumentar la memoria de trabajo. Un ser sobrehumano tendría una memoria ilimitada para los resultados obtenidos durante una computación, como por ejemplo los productos intermedios que necesitamos recordar si tratamos de hacer mentalmente multiplicaciones largas. Por tanto, a medida que la capacidad de memoria de trabajo de los niños aumenta, son capaces de procesos más complejos. Como ejemplo consideremos otra tarea piagetiana en la que los niños tienen que ordenar un conjunto de varillas de acuerdo con su longitud. Los niños de cuatro años no pueden hacer esta tarea. Sus distribuciones suponen solo una burda aproximación al orden correcto de la serie. Pueden iniciar una serie con una

varilla larga y colocar la siguiente varilla que cogen en el lugar correcto. Pero a partir de entonces actúan a voleo, y tienen dificultades para interpolar una nueva varilla entre las que ya están colocadas. La interpolación les exige razonar sobre las relaciones. Necesitan razonar que debido a que la nueva varilla es más larga que una de las varillas y más corta que la otra, debería colocarse entre ellas. Esta inferencia es imposible si nuestra memoria de trabajo no puede contener a la vez más de una relación binaria entre dos objetos. A los cinco años, sin embargo, la capacidad de memoria de trabajo de los niños ha aumentado, y pueden colocar las varillas en el orden correcto, aunque siguen teniendo dificultades para interpolarlas. En lugar de interpolarlas tienden a recolocar pares de varillas contiguas, y la recolocación se transmite a toda la serie hasta que está en el orden correcto. Aproximadamente a los siete años, su memoria de trabajo es todavía mayor, y pueden hacer una colocación sistemática de las varillas, y enfrentarse a las interpolaciones.

Los psicólogos miden la capacidad de la memoria de trabajo de varias formas. La forma más simple es medir cuántos dígitos puede retener un niño durante algunos segundos. Usted y yo podemos hacer frente a un número de teléfono de unos siete dígitos, a veces más, a veces menos. Un niño de tres años, sin embargo, puede hacer frente solo a unos tres dígitos. A los cuatro años, un niño puede hacer frente a cuatro dígitos. Esta mejora en la memoria de trabajo parece ser una consecuencia de la maduración del cerebro, pero también una consecuencia del desarrollo del conocimiento. Los niños adquieren formas más eficaces de representar información en la memoria de trabajo. A continuación paso a un aspecto diferente de la arquitectura mental.

* * *

La evolución no puede haber dado forma a una mente que dependa de una masa homogénea. Debe haber estado reciclando un amplio conjunto de “órganos” mentales independientes, y por tanto la mente puede dividirse en módulos independientes, más o menos por la misma razón por la que los programadores desarrollan sus programas en rutinas independientes. Los programadores quieren trabajar sobre una rutina sin tener que pensar en todas al mismo tiempo. Por tanto, una división en rutinas independientes es útil en programación. De forma

similar, si el cerebro fuera una única masa de neuronas sin diferenciar, la evolución no podría reciclarlo de una forma que le permitiera adaptarse. El reciclado puede ser adaptativo debido a que afecta a uno de los módulos pero no a todos. Una variación en un módulo para la percepción del color podría conducir a una mejor actuación sin tener efectos adversos en ningún otro módulo. Los individuos con esta variación tendrían una mejor oportunidad de sobrevivir y reproducirse, y por tanto sus genes se harían más frecuentes en la población.

Un reciente grupo de teóricos, los "psicólogos evolucionistas", ha ido un paso más allá. Imaginan qué es lo que habría dado esa capacidad de adaptarse a nuestros ancestros cazadores-recolectores hace unos 120.000 años, postulan un módulo innato para computarlo, y luego llevan a cabo experimentos para ver si el módulo existe todavía en nuestro pensamiento. Un módulo innato sirve, dicen, para razonar sobre si alguien nos está engañando. Habría surgido debido a que los intercambios sociales habrían sido importantes para nuestros ancestros. Otro módulo sirve, dicen, para razonar sobre precauciones frente a los riesgos.

El razonamiento tiene realmente una estructura modular. Cada uno de los procedimientos que describí para el razonamiento con modelos es un módulo independiente en el programa de simulación de la teoría (ver la Tabla 9.2). Por ejemplo, hay un módulo para empezar la construcción de un modelo, uno para actualizarlo, etc. Y este tipo de módulos se pueden utilizar de diferentes maneras para la deducción, la inducción y la abducción. ¿Pero hay módulos independientes para razonar sobre diferentes temas? Es decir, ¿procesos que se diferencien dependiendo de si estamos razonando, pongamos por caso, sobre personas que pueden estarnos engañando o sobre cómo evitar contraer el cólera? Sabemos que el contenido tiene importancia. Por ejemplo, describí anteriormente cómo afecta el contenido a la tarea de selección (en el Capítulo 1). Y por tanto se podría suponer que debe haber módulos diferentes para contenidos diferentes. Una hipótesis alternativa, sin embargo, es que un único proceso de razonamiento lleve a cabo diferentes operaciones dependiendo de las diferencias en el contenido. El contenido puede bloquear la construcción de posibilidades, añadir información a un modelo y traer a la mente un contraejemplo conocido.

¿Hay un módulo innato para razonar sobre el engaño? Quizá no, porque la noción de engaño es vaga. ¿Es engaño mentir? ¿Es engaño la poligamia? ¿Es engaño el asesinato de un enemigo? El problema para la teoría es explicar qué

desencadena la utilización de un módulo apropiado. Los psicólogos evolucionistas necesitan enumerar de forma exhaustiva los módulos innatos, una explicación de qué es lo que desencadena cada uno de ellos y una descripción de las diferentes formas en que funcionan. El respaldo a la teoría evolucionista del razonamiento depende en la actualidad de la tarea de selección. Pero los críticos señalan que los resultados se pueden explicar de otras formas. Y el contenido no siempre afecta al razonamiento. Varios estudios no han conseguido encontrar ninguna diferencia significativa entre el razonamiento sobre hechos y el razonamiento sobre obligaciones. No existen pruebas decisivas que exijan la existencia de módulos innatos que razonen sobre contenidos diferentes. Los seres humanos tienen solo unos treinta mil genes, y el utilizarlos en la codificación de módulos para razonar sobre temas concretos parecería un considerable despilfarro.

Sin duda la evolución ha dado forma a la mente humana. Pero es difícil decir por qué vías, más allá de módulos como los del oído o la visión, que evolucionaron mucho antes de que lo hicieran los seres humanos. Como dicen los críticos de la psicología evolucionista, no podemos volver a la Edad de Piedra para evaluar la selección natural mientras está actuando. Lo mejor que podemos hacer es contar una historia “simple” sobre lo que habría dado capacidad de adaptación a nuestros ancestros, y evaluar si predice cómo trabaja hoy nuestra mente. Los psicólogos evolucionistas sí que inventan algunas historias divertidas – desde cómo consiguieron los humanos su gusto por el papel pintado floreado hasta por qué se pelean los hermanos–. Pero, ¿qué deberían hacer cuando una historia simple no lo es tanto? Pueden admitir –y lo hacen– que no todas las funciones mentales son adaptaciones, o pueden reescribir la historia para que concuerde con los hechos. La psicología evolucionista es un heurístico útil para generar hipótesis. Estas hipótesis se pueden evaluar, pero es difícil ver cómo evaluar la teoría en conjunto.

* * *

Una gran parte del razonamiento se puede haber desarrollado a partir de capacidades innatas. Estas capacidades subyacen en la construcción de modelos a partir de la percepción, la imaginación y el lenguaje. De ahí que los bebés se

centren en la violaciones de la física ingenua, y los niños puedan razonar sobre lo que hay en la mente de los demás. Quizá, todos los mecanismos fundamentales del razonamiento sean innatos. Sin embargo, la capacidad de razonamiento de un bebé recién nacido es muy diferente de la de un Einstein maduro. Por tanto, ¿cómo pasamos de una a otra?

Un factor es el crecimiento del conocimiento. Por una sutil ironía, las pruebas proceden de una línea de investigación que intentaba mostrar algo diferente, a saber, los estudios de Piaget sobre la conservación. El dominio de las operaciones reversibles, decía Piaget, depende de una reestructuración general de la mente que se produce en torno a los siete años (véase la Tabla 18.1). Pero una posibilidad alternativa es que los niños no adquieran un principio general para las operaciones reversibles, sino el conocimiento específico de que deformar plastilina, verter agua o recolocar vasos no afecta a la cantidad. Los invariantes no son necesidades lógicas. Los niños tienen que aprenderlos de uno en uno, lo mismo que tienen que aprender que cuando el agua se congela aumenta de volumen. Como ejemplo, considere esta demostración. Cojo un lazo de cuerda y lo sostengo sobre mis pulgares y mis índices para darle forma de cuadrado. A continuación, manteniendo tensa la cuerda en todo momento, alejo las manos pero acerco los pulgares y los índices de modo que la forma de la cuerda se convierta en un rectángulo. ¿Sigue la cuerda encerrando el mismo área? La mayoría de la gente dice, “sí”. La conservación es la norma. Pero la respuesta correcta es: “no”. Puede hacerlo utilizando la aritmética, o puede imaginar que extiendo la cuerda de manera que la altura del rectángulo se haga tan pequeña que sea invisible, y en este caso el área del rectángulo debe ser menor. Si no sabe usted nada sobre lazos de cuerda, no se necesita ninguna reestructuración mental general para adquirir este conocimiento, solo la adquisición específica de una pequeña verdad sobre el mundo. Otra prueba en favor de esta visión específica es que cuando el conocimiento tiene importancia los niños lo adquieren antes. Los hijos de los aborígenes australianos que viven en las zonas del interior del país no pueden correr el riesgo de perderse, y aprenden a conservar las relaciones espaciales entre distintas perspectivas de un paisaje antes que los que viven en la ciudad.

Los psicólogos han investigado cómo se desarrolla el conocimiento. Inicialmente los niños establecen categorías de objetos a partir de rasgos superficiales. Si anda como un pato y habla como un pato, entonces es un pato. Más tarde, empiezan a captar la importancia de factores causales ocultos. Si es la cría de un pato, entonces es un pato aunque no tenga su forma de andar o de hablar. De

manera similar, sus concepciones de la tierra y el sistema solar parten de modelos supersimplificados, o fragmentos de conocimiento, y avanzan hacia concepciones más maduras.

El cómo pasan los niños desde un modelo del mundo a otro no está muy claro y hasta ahora ningún programa de ordenador lo ha simulado. La psicóloga evolutiva Susan Carey asemeja el proceso con el derrocamiento de una teoría científica por otra. Quienes proponen una teoría tienen dificultades para entender la teoría rival. No les resulta imposible hacerlo, pero es un trabajo duro, porque las teorías parecen no poderse comparar. Por ejemplo, Goethe tuvo problemas para entender la teoría del color de Newton. El objetivo de Newton era comprobar su teoría mediante la descomposición de la luz en los colores del espectro y mediante la mezcla de colores al hacer girar ruedas con secciones de colores diferentes. En cambio, los experimentos de Goethe tenían carácter exploratorio y estaban más cerca de la psicología. Newton consideraba los colores del espectro como fundamentales y Goethe permitía que colores que no aparecen en el espectro, como el magenta, también fueran fundamentales. Carey ha señalado una imposibilidad similar para comparar los conceptos de los niños pequeños y los conceptos de los adultos. Informa de este diálogo con su hija de tres años y medio:

Eliza: ¿Cómo van al baño las personas muertas?

Carey: ¿Qué?

Eliza: ¿Quizá tengan baños bajo el suelo?

Carey: Las personas muertas no van al baño. No hacen nada; solo están allí. No comen ni beben, por eso no necesitan ir al baño.

Eliza: Pero comieron y bebieron antes de morir - [triunfante] ¡necesitan ir al baño desde antes de que se murieran!

Eliza pensaba que los muertos viven bajo tierra en condiciones diferentes. No es más fácil sacar a los niños de sus concepciones erróneas que sacar a los científicos de sus teorías favoritas.

* * *

Los que no tienen una preparación científica cargan con su “física ingenua” desde la infancia a la madurez. Estudios realizados en Italia y posteriormente en Estados Unidos mostraron una sorprendente similitud entre las concepciones erróneas de los adultos respecto a la física y las opiniones de Aristóteles y las de sus seguidores medievales. Aristóteles pensaba que el movimiento era un cambio de lugar. Se producían dos tipos de movimiento: movimientos naturales, como la caída de objetos pesados a la Tierra, y movimientos no naturales, como el movimiento de un objeto que es empujado por otro. La idea de rapidez de Aristóteles era que de dos objetos en movimiento, el más rápido llega más lejos en el mismo tiempo, y recorre la misma distancia en menos tiempo. Estas ideas son similares a las de sentido común. El movimiento es un cambio de posición en el tiempo. Algunos objetos se mueven de manera natural; otros se mueven debido a la acción de otros objetos sobre ellos. Tendemos a pensar que un objeto al que se le ha dado un empujón tiene un “impetus” que se agota, y por tanto el objeto se frena y al final se para. Asimismo, consideremos lo que pensamos de un objeto que está girando en círculo al extremo de una cuerda cuando la cuerda se rompe. Muchos imaginamos que el objeto se mantiene en una trayectoria espiral que se abre lentamente hasta que acaba viajando en línea recta. Suponemos que el “impetus” circular se desenrolla. Cometemos un error similar respecto a la trayectoria de una bola que sale de un tubo con forma de C. Sin embargo, nos abstenemos de cometer este error con un chorro de agua, quizá porque hemos observado el agua que sale de un grifo curvado. De acuerdo con la mecánica newtoniana, una fuerza centrípeta acelera al objeto de la cuerda o del tubo curvado hacia su centro de rotación. Cuando la cuerda se rompe o el objeto sale del tubo, la aceleración cesa. Se aplica entonces la primera ley del movimiento: un objeto permanece en reposo o se mueve con velocidad uniforme en línea recta si no existe una fuerza que actúe sobre él. Por tanto, el objeto escapa en línea recta.

Nuestra concepción de sentido común respecto a la rapidez también nos confunde. Como Aristóteles, concebimos que un objeto se mueve más rápido que otro si en un determinado momento están alineados y en un momento posterior un objeto está por delante del otro. Pero también podemos decir que un

objeto se mueve con más velocidad si parece borroso. Cuanto más borroso, mayor es la rapidez. Nuestros modelos mentales plasman ambas nociones: un objeto adelanta a otro (tiene una mayor rapidez media), y se mueve más rápido porque está más borroso (rapidez instantánea).

Las revisiones científicas de estas concepciones llevaron varios siglos. La mecánica newtoniana se aplica a los dos tipos de movimientos de Aristóteles, el natural y el no natural. El sentido común establece una distinción innecesaria. Asimismo, la concepción de sentido común respecto a la rapidez es incoherente. Su carácter confuso se ilustra en este diálogo adaptado de las clases de física del difunto Richard Feynman:

Policía: Iba usted a 60 millas por hora.

Conductor: Es imposible. Solo he viajado durante siete minutos. ¿Cómo puedo recorrer 60 millas por hora, cuando no he viajado durante una hora?

Policía: Lo que quiero decir es que si usted siguiera yendo como iba, entonces en la siguiente hora recorrería 60 millas.

Conductor: El coche estaba frenando, por tanto, si siguiera yendo así no recorrería 60 millas. Y en cualquier caso, ¡si hubiera seguido como iba, habría chocado con el muro del final de la calle!

Policía: Lo que quiero decir es que si usted anduviera un segundo recorrería 88 pies.

Conductor: No existe una ley en contra de recorrer 88 pies por segundo.

Policía: ¡Dígale eso al juez!

El policía podría haber replicado en lugar de eso que 88 pies por segundo es lo mismo que 60 millas por hora. Así que la rapidez del conductor en el momento en que el destello del radar se reflejó en la parte trasera del coche era 60 millas por hora. Este concepto de rapidez instantánea debe distinguirse de nuestra noción de sentido común respecto a la rapidez media. La rapidez instantánea se

basa en el concepto matemático de “diferenciación” debido a Newton, es decir, dividimos la distancia por el tiempo mientras hacemos que el intervalo de tiempo sea más y más pequeño, y observamos el resultado cuando el intervalo tiende a ser infinitesimal. En este caso, una única noción incipiente se perfecciona para dar dos conceptos distintos: rapidez instantánea y rapidez media.

En la teoría de Newton, la masa de un objeto que se mueve es constante, siendo la masa la inercia de un objeto, es decir, lo difícil que resulta poner el objeto en movimiento aplicándole una fuerza. Pero, en la teoría especial de la relatividad de Einstein, la masa de un objeto ya no es constante, sino que aumenta con su velocidad. A medida que esa velocidad se aproxima a la rapidez de la luz, la masa del objeto se aproxima a infinito. Por ello, no puede ir más rápido que la luz. Este nuevo concepto coincide parcialmente con el antiguo. Comparten el concepto de inercia, sin embargo, difieren en cómo le afecta el movimiento. Pero estos conceptos corresponden a físicos expertos, y no voy a ir más allá en relación con ellos.

Una revolución en ciencia conduce a un nuevo tipo de modelo, que con frecuencia incorpora conceptos que no se pueden observar en el mundo. Puede llevar a varios tipos de cambio conceptual: la mezcla de dos nociones independientes, o la división de un único concepto incipiente en dos conceptos independientes, o el mantenimiento de algunos componentes de un concepto y la modificación de otros. Nuestro desarrollo intelectual es similar, pero a menos que tengamos una educación científica, muchos de nuestros conceptos siguen siendo incoherentes.

* * *

Si no conocemos los significados de las conectivas o los cuantificadores, nos equivocaremos al hacer inferencias basadas en ellos. Podemos no darnos cuenta de que una oración como por ejemplo:

Todos los artistas son apicultores

no significa que:

Todos los apicultores son artistas.

Podemos adoptar la hipótesis más simple, a saber, que quienes son artistas son apicultores, y viceversa. Nuestra hipótesis es útil porque minimiza el número de referentes distintos que tenemos que tener presentes. Pero nos conduce al camino equivocado. Inferimos que todos los apicultores son artistas porque todos los artistas son apicultores. Nos equivocamos por nuestra falta de conocimiento sobre el significado de los cuantificadores. Asimismo, si no tenemos el conocimiento que nos capacitaría para modular la interpretación de un término, tampoco conseguiremos extraer inferencias correctas.

Un aumento de la capacidad de la memoria de trabajo da lugar también a un mejor razonamiento. El psicólogo francés Pierre Barrouillet y sus colegas han mostrado que cuando a los niños se les pide que enumeren las posibilidades compatibles con un condicional como por ejemplo:

Si hay un triángulo, entonces hay un círculo

tienden a nombrar solo una posibilidad en la que hay ambas cosas, un triángulo y un círculo. Tratan el condicional como si fuera una conjunción. Al hacerse mayores, enumeran dos posibilidades: una en la que hay ambas cosas, el triángulo y el círculo, y una en la que no hay ni triángulo ni círculo. Y, cuando se aproximan a la madurez, enumeran tres posibilidades compatibles con el significado “lógico” del condicional: añaden el caso en el que no hay un triángulo pero hay un círculo. Lo que mejor predice el número de posibilidades en que piensan los niños no es la edad, sino la capacidad de su memoria de trabajo.

Los efectos de la memoria de trabajo también se producen cuando los niños extraen sus propias conclusiones a partir de las premisas, como por ejemplo:

Algunos de los padres son profesores.

Todos los profesores son conductores.

Esta inferencia exige la construcción de un modelo mental único, y los niños de solo siete años pueden extraer la conclusión:

Algunos de los padres son conductores.

En contra de la teoría de Piaget, se enfrentan a estas inferencias con mejores resultados de los que corresponderían a la casualidad. Sin embargo, cuando las inferencias exigen modelos múltiples, la mayoría de nosotros resultamos derrotados. La capacidad para hacer inferencias mejora con la edad entre los siete años y la edad adulta, y esta mejora correlaciona con un aumento en la capacidad de la memoria de trabajo.

* * *

El desarrollo del lenguaje depende tanto de la naturaleza como de la educación. Los bebés de pocos días ya se han acostumbrado a la lengua de su madre y atienden más a ese que a otros lenguajes. Pero, si crecen sin poder oír nunca un lenguaje o ver un lenguaje de signos, entonces, como ocurre en el reducido número de los trágicos casos de “niños salvajes”, nunca dominan un lenguaje.

Por tanto, tanto la dotación genética como el entorno, son importantes. Ambos interaccionan. El razonamiento puede depender también de una base innata para imaginar posibilidades, y para manipular sus modelos. Su desarrollo, como he sostenido, se apoya en el aumento del conocimiento, incluidos conceptos y significados, y en un aumento de la capacidad de la memoria de trabajo. El cómo aumenta el conocimiento es desconcertante. No es la mera adición de hechos. Exige revisiones de los conceptos, y quizá cambios en nuestros modelos de la realidad. El proceso es un misterio, y actualmente supera la capacidad de explicación de los científicos cognitivos.

CAPÍTULO 19

Estrategias y culturas

Hace algunos años un distinguido científico cognitivo vino a Princeton a dar una charla. Llegó antes de tiempo para charlar conmigo, pero yo no estaba en mi despacho, y para entretenerle un estudiante de postdoctorado le planteó una de las inferencias ilusorias en las que estábamos trabajando. La mayoría de la gente la hacía mal, y él también la hizo mal. Más tarde, me explicaba:

Yo me dije: “Este es uno de los tontos problemas de inferencias de Phil, y por tanto tengo que tener cuidado.” ¡Pero aun así lo hice mal!

Este comentario ilustra el razonamiento sobre razonamiento –al que me referí anteriormente como “metarrazonamiento”–. A pesar de su insight, no consiguió llegar a una estrategia adecuada para tratar la inferencia. En este capítulo, voy a considerar un tipo especial de desarrollo, el desarrollo a corto plazo de estrategias de razonamiento, que se produce cuando tratamos de resolver una serie novedosa de problemas de razonamiento. Una estrategia es una secuencia de pasos que nos capacitan para inferir la respuesta a un problema deductivo. En capítulos anteriores, cuando me he referido a un método o procedimiento para hacer una inferencia, estaba hablando de una estrategia. Al pensar en voz alta mientras hacemos una inferencia, describimos cada paso de nuestra estrategia, y cada uno de esos pasos es lo que llamaré una táctica. Somos conscientes de nuestras tácticas, y de la secuencia según la cual se producen, pero no somos conscientes de los mecanismos que subyacen en ellas. Yo sostengo que estos mecanismos están enraizados en modelos mentales, pero incluyen mecanismos en relación con la gramática y con la detección de regularidades formales.

En la vida real, razonamos continuamente, pero nuestras inferencias son simples, están insertadas en un contexto razonable, invocan nuestro conocimiento, y no

implican hacer pedazos la lógica. Nos enteramos, por ejemplo, de que la vacuna de la gripe estará disponible solo para individuos de riesgo, como los niños y los ancianos. Sabemos que no estamos en riesgo, y por tanto nos damos cuenta –inferimos, para ser preciso– que no recibiremos un pinchazo. Nuestras estrategias son simples a menos que hayamos aprendido a hacer frente a inferencias de un cierto tipo, tales como el diagnóstico médico, arreglar radios estropeadas, o la ingeniería inversa de las tostadoras. En los experimentos, se nos presentan una serie de problemas que están estandarizados, aislados de cualquier contexto cotidiano y desconectados del conocimiento general. Cuando nos encontramos con una inferencia de este tipo por primera vez, por un momento nos sentimos desconcertados. Pero tendemos a ser capaces solucionarla. Mientras vamos realizando unas cuantas inferencias de ese tipo, desarrollamos una estrategia para enfrentarnos a ellas. Pero, como veremos, diferimos unos de otros respecto a cuáles son las que adquirimos. Un corolario obvio es que no llegamos al laboratorio previamente equipados con una estrategia para hacer frente a estas inferencias. A medida que la desarrollamos, nuestra actuación mejora. Aprendemos a razonar mejor.

En el pasado, los psicólogos han investigado las estrategias que se utilizan al razonar, pero utilizando problemas basados en premisas sencillas sobre relaciones, como por ejemplo: Pat es más alta que Viv. Cuando los participantes tratan de resolver problemas basados en cinco premisas de este tipo, enseguida desarrollan atajos. Por ejemplo, dado un problema con premisas que siempre contienen la misma relación, como:

Beth es más alta que Cath.

Ann es más alta que Beth.

Dot es más alta que Eve.

Cath es más alta que Dot.

Eve es más alta que Fran.

¿Quién es la más alta?

los participantes buscan si un término aparece solo en la parte izquierda de una única premisa. Indica el individuo más alto –Ann en el caso considerado–. Cuando en el problema se plantea una pregunta, como por ejemplo, ¿Cath es más alta que Beth?, los participantes pueden utilizarla para seleccionar solamente aquellas premisas que necesitan para construir una cadena que enlace a esos dos individuos. Y, cuando una inferencia tiene dos premisas, algunas veces tratan de resolver la cuestión solamente a partir de la información de la primera premisa. A veces funciona, pero puede salir mal.

Algunos psicólogos sostienen que nuestro razonamiento deductivo depende de una estrategia única. Pero, a lo largo de los años, los experimentos han descubierto varios problemas respecto a esta opinión. El orden de las premisas, por ejemplo, tiene fuertes efectos sobre las inferencias a partir de premisas condicionales. Estos efectos parecen no ser consistentes con una estrategia única. Y la hipótesis parecía haber sobrevivido porque los experimentos no eran lo suficientemente sensibles como para detectar las estrategias. No utilizaban más de dos premisas, solamente examinaban las conclusiones que extraían los participantes, y algunas veces la rapidez de sus respuestas. Estos resultados no podían revelar cómo llegaban los participantes a sus conclusiones, y por tanto los psicólogos no prestaban suficiente atención a la posibilidad de que los participantes desarrollaran diversas estrategias.

Mi intención es corregir esa falta de atención. Y por tanto, voy a describir algunos estudios que revelan la existencia de estrategias en diversos tipos de razonamiento. Estos estudios corroboran la existencia de cuatro “niveles” diferentes de razonamiento. En el nivel más alto está el metarrazonamiento que se pone de manifiesto en los comentarios de mi visitante. En el nivel más bajo están los mecanismos inconscientes que describí en la Parte II de este libro. Lo que me importa ahora son los dos niveles intermedios: las estrategias y las tácticas que las componen. El psicólogo Robert Sternberg fue pionero en el

estudio de esas tácticas en sus estudios sobre analogías, problemas de series numéricas y silogismos. Posteriormente, en este mismo capítulo, sostengo que las estrategias son fundamentales en el debate sobre las diferencias en el razonamiento entre una cultura y otra.

* * *

Una reciente moda universal que comenzó en Japón es la de los Sudokus –el nombre es una abreviatura japonesa de “un único número”–. Los rompecabezas se introdujeron en Gran Bretaña en noviembre de 2004 en el Times de Londres. Se extendieron a muchos periódicos británicos nacionales y luego a otros países de Europa, a Canadá y Australia, y al New York Post en julio de 2005. Me atrevería a decir que en el momento en que usted está leyendo este párrafo están por todas partes o, tal como van las modas, en ninguna. Son maravillosos para quienes estudian el razonamiento porque son los primeros rompecabezas populares que dependen de la pura deducción. Son los “Rompecabezas de hielo de la razón” de Hans Andersen por cuya solución la Reina de las Nieves prometió dar al pequeño Kay “el mundo entero y un par de patines”.

La Figura 19.1 es un ejemplo de un Sudoku. Por si usted no se hubiera encontrado nunca con uno, le explicaré cómo funciona. Aunque contiene números no exigen ningún conocimiento aritmético –las letras del alfabeto funcionarían igual–. La tarea es rellenar los números que faltan en el esquema de acuerdo con tres reglas:

7		1			3		6	5
		4	6				3	
			5	1	9			
	4	3	9			6		7
6				5				1
2		9			7	8	5	
			1	9	6			
	3				8	4		
1	8		4			9		2

Fig. 19.1 Un Sudoku fácil (cuya solución aparece en las notas al capítulo).

Cada fila horizontal del rompecabezas debe contener solo una vez cada uno de los dígitos del 1 al 9.

Cada columna vertical del rompecabezas debe contener solo una vez cada uno de los dígitos del 1 al 9.

Cada recuadro de tres por tres del rompecabezas –hay nueve recuadros de ese tipo destacados en negrita– debe contener solo una vez cada uno de los dígitos del 1 al 9.

Cada una de estas reglas depende de tres cuantificadores: cada, solo una y cada uno: son reglas que contienen múltiples cuantificadores.

Los Sudokus son experimentos naturales sobre razonamiento y establecen tres hechos. Ítem: muchos de nosotros, incluso aunque no hayamos recibido formación en lógica, somos capaces de hacer deducciones que dependen de múltiples cuantificadores. Disfrutamos enfrentándonos a deducciones de este tipo. Este hecho es contrario a la opinión de que sin una formación lógica somos incapaces de hacer deducciones válidas basadas en cuantificadores, y a la opinión de que la validez es un criterio inadecuado para nuestras deducciones. Ítem: la capacidad deductiva es común a todas las poblaciones del mundo. Esto es contrario a la hipótesis de que las deducciones son una prerrogativa de los occidentales, mientras que los orientales las rechazan en favor del pensamiento "holístico". Ítem: cuando tratamos de resolver una serie de Sudokus, desarrollamos estrategias para resolverlos basadas en la adquisición de tácticas deductivas particulares, y, como resultado, aprendemos cómo resolver problemas más difíciles. Esta observación es contraria a la tesis de que todos razonamos utilizando la misma estrategia fija, y también es contraria a la opinión de que

nuestra capacidad de razonamiento no puede mejorar.

Los Sudokus aparecen en cinco niveles de dificultad: triviales, que nunca se publican pero son sencillos de realizar, fáciles, medios, difíciles y extremadamente difíciles (o, como le llaman los periódicos italianos, “diabolico”). Estos niveles los establecen los creadores de los Sudokus, pero saben de lo que están hablando. Para alguien competente en su resolución, un Sudoku fácil puede resolverse en unos diez minutos, mientras que uno muy difícil puede llevar casi dos horas. Parte de la atracción de estos rompecabezas puede ser esta gran diferencia de dificultad, ya que a medida que dominamos los más sencillos podemos pasar a otros más difíciles. Cuando resolvemos nuestro primer Sudoku, como confirmó un experimento, tendemos a utilizar tácticas sencillas. Estas tácticas nos permiten establecer la identidad segura de un dígito en un solo paso. Por ejemplo, cuando en una línea o en un cuadro se han completado todos los dígitos menos uno, entonces el valor de la casilla que falta es el número que falte entre el 1 y el 9. Rápidamente adquirimos otras tácticas sencillas. Consideremos, por ejemplo, el cuadro de 3 x 3 de la izquierda en la fila superior de la Figura 19.1. Contiene ya 1, 4 y 7, y la regla general de los recuadros implica que debe haber un 3 en alguna parte de ese recuadro. Ya hay un 3 en las dos filas superiores que se cruzan con él, y por tanto el 3 debe aparecer en la fila inferior del recuadro. Pero los números 3 en las columnas de debajo del recuadro excluyen todas las posibles casillas excepto la esquina inferior izquierda. Por tanto, como resultado de estas cinco limitaciones independientes (el propio recuadro, las dos filas y las dos columnas) el 3 debe estar en esa casilla. Las tácticas sencillas difieren en dificultad como consecuencia del número de limitaciones que tenemos que tener en cuenta, que varían de una a cinco. Los psicólogos se refieren al número de limitaciones de una inferencia como a su “complejidad relacional” y un experimento mostró que los participantes tendían a necesitar más tiempo para utilizar tácticas sencillas de Sudoku a medida que el número de limitaciones aumentaba, aunque las cinco limitaciones eran un poco más fáciles que las cuatro, quizá porque cuatro ejemplos del mismo dígito (en el caso de cinco limitaciones) son más fáciles de descubrir en la distribución que tres dígitos (en el caso de cuatro limitaciones).

Las tácticas sencillas nos capacitan para progresar en los Sudokus de todos los niveles de dificultad, y bastan para resolver la mayoría de los problemas fáciles y medios. Sin embargo, tenemos que hacer un importante cambio en la estrategia con el fin de resolver los Sudokus difíciles y muy difíciles. Tenemos que aprender a enumerar los dígitos posibles en las casillas. Solo cuando lo hacemos

podemos desarrollar tácticas avanzadas. Nos capacitan para eliminar valores que son posibles para las casillas. Por ejemplo, supongamos que un par de casillas de una línea o cuadro tienen solo dos posibles valores, p. ej., 4 u 8. Una casilla debe contener el 4, y la otra casilla debe contener el 8, aunque no sabemos qué casilla es cada una. Se sigue entonces que estos dos dígitos se pueden eliminar como posibilidades de todas las demás casillas de la misma línea o recuadro.

Podría usted pensar que la diferencia de dificultad entre los Sudokus depende del número de casillas vacías en la disposición inicial. Existe una ligera tendencia a que en los Sudokus más difíciles el número de dígitos eliminados sea mayor, pero esto en sí mismo no es una causa fundamental de dificultad. Sudokus en los que falten el mismo número de dígitos pueden variar mucho en cuanto a su dificultad. Incluso cuando se han completado muchos dígitos en un Sudoku muy difícil, se puede llegar a una situación casi imposible. Y, al empezar a solucionar un rompecabezas, los pasos tácticos iniciales a menudo no difieren en dificultad de un nivel de a otro.

Si un rompecabezas está en un determinado estado en cuanto a su solución, se han resuelto un cierto número de dígitos, y se pueden hacer un cierto número de deducciones independientes unas de otras respecto a las casillas restantes. Algunas de estas deducciones pueden hacernos determinar el dígito seguro que corresponde a las casillas, y otras pueden eliminar posibles valores de los dígitos de las casillas. Un estado finaliza cuando se han hecho todas esas deducciones para actualizar las posibilidades en la distribución. A continuación, en el siguiente estado, puede hacerse una nueva serie de deducciones independientes, teniendo en cuenta las deducciones previas, etc. No tendemos a hacer todas las deducciones que se pueden inferir en cada estado. Algunos hacemos algunas deducciones, otros hacemos otras, y la secuencia varía de una persona a otra al menos hasta los estados finales. No obstante, cuanto más fácil es un rompecabezas mayor debería ser por término medio el número de dígitos seguros que se pueden deducir en cualquier estado. Escribí un programa de ordenador para resolver Sudokus y lo utilicé para comprobar esta predicción. Tomé una muestra aleatoria de diez rompecabezas de cada uno de los cuatro niveles de dificultad (fácil, medio, difícil y extremadamente difícil) sacados de un compendio de Sudokus. El programa estableció el número medio de dígitos seguros que se podían inferir en cada estado al solucionar los problemas y los resultados eran 4,3 para los rompecabezas fáciles, 2,3 para los rompecabezas medios, 1,8 para los rompecabezas difíciles y 1,3 para los rompecabezas extremadamente difíciles. Era totalmente improbable que se diera esa tendencia

por casualidad. Por tanto, para cada estado de un rompecabezas, se pueden inferir más números seguros en los rompecabezas más fáciles que en los más difíciles. Como corolario, si en cada estado hacemos todas las deducciones posibles, entonces los rompecabezas más sencillos necesitan por término medio menos estados que los rompecabezas medios, etc. Sin embargo, las tácticas sencillas por sí solas no pueden resolver los rompecabezas difíciles o extremadamente difíciles. Los Sudokus son una demostración clara de que desarrollamos estrategias de razonamiento, que a su vez nos capacitan para mejorar nuestra capacidad para deducir sus soluciones. Pero, para seguir la pista a cómo desarrollamos estrategias, voy a considerar deducciones mucho más sencillas.

* * *

La Tabla 19.1 representa un ejemplo típico del tipo de problemas que se utilizan en los experimentos sobre razonamiento deductivo, excepto porque frente a las dos o tres premisas de las que son partidarios los psicólogos, este tiene algunas más. Si, como imagino, no se ha molestado en resolver el problema, por favor resuélvalo ahora. Es fácil. Puede utilizar lápiz y papel, pero trate de pensar en voz alta a medida que razona –para que más tarde tenga alguna idea sobre su propia estrategia.

Tabla 19.1. Un problema típico utilizado en un estudio sobre estrategias de razonamiento

Dentro de la caja hay una canica roja si y solo si dentro de la caja hay una canica marrón.

O bien dentro de la caja hay una canica marrón o bien dentro de la caja hay una

canica gris, pero no ambas.

Dentro de la caja hay una canica gris si y solo si dentro de la caja hay una canica negra.

¿Se sigue que: Si no hay una canica roja dentro de la caja entonces dentro de la caja hay una canica negra?

(Y bien, ¿se sigue? Intente pensar en voz alta mientras resuelve este problema.)

La respuesta correcta al problema es: sí, si dentro de la caja no hay una canica roja entonces hay una canica negra. Un experimento basado en problemas de este tipo examinaba las estrategias que desarrollaban por sí mismos los participantes. Las premisas de cada problema eran compatibles con solo dos posibilidades, y la mitad de los problemas se presentaban con una conclusión válida y la otra mitad se presentaban con una conclusión no válida. Los participantes tenían que pensar en voz alta a medida que razonaban, y se les permitía utilizar lápiz y papel. Sobre ellos y enfocada a la mesa en la que estaban escribiendo se situaba una cámara de vídeo. Se adaptaban rápidamente a su presencia. La introspección puede ser una guía poco fidedigna para nuestro pensamiento, porque podemos racionalizar. Pero, cuando los participantes pensaban en voz alta mientras estaban razonando, lo que decían sí que revelaba sus principales pasos –la secuencia de sus tácticas– a partir de los cuales es posible reconstruir sus estrategias. Sin embargo, nunca decían nada sobre sus estrategias, nada del tipo de “mi intención es construir un diagrama sencillo basado en todas las premisas a partir del que puedo comprobar si se sigue la conclusión”. Y, como podría esperarse, no decían nada sobre el mecanismo que subyacía en sus pasos tácticos –no podían porque los mecanismos son inconscientes.

Todos los participantes hicieron correctamente todos los problemas, y sus

estrategias eran claras para casi todos los problemas, aunque algunas veces cambiaban de una estrategia a otra en mitad de alguno. Este experimento, y los posteriores, mostraron que los participantes tienden a desarrollar cinco estrategias principales para el razonamiento basado en conectivas proposicionales, del tipo de la inferencia de la Tabla 19.1. La Tabla 19.2 resume las cinco estrategias. La mayoría de los participantes desarrollaron una o dos durante una sesión experimental.

Tabla 19.2. Cinco estrategias para el razonamiento basado en las conectivas proposicionales

1. Diagrama único: Dibujar un diagrama que sigue el rastro de las posibilidades independientes consistentes con las premisas.

2. Posibilidad única: Seguir paso a paso las consecuencias de una única posibilidad. Si un problema incluye una premisa que hace una afirmación categórica, entonces seguir sus consecuencias. En otro caso, hacer una suposición para iniciar la puesta en marcha de la estrategia, o bien una sugerida por una premisa, o quizá una que sea un contraejemplo de una conclusión dada.

3. Cadena: construir una cadena de condicionales que lleven de una cláusula de una conclusión condicional a su otra cláusula, convirtiendo cualquier premisa en un condicional adecuado de forma que conduzca de un condicional de la cadena al siguiente.

4. Compuesta: diseñar una conclusión compuesta a partir de un par de premisas compuestas, en las que “compuesta” significa que un aserto contiene una conectiva. Combinar la conclusión con una nueva premisa compuesta, y así sucesivamente, hasta que se alcanza la conclusión.

5. Concatenación: concatenar dos o más premisas con el fin de formar una conclusión, p. ej., premisas de la forma: A si y solo si B; O bien B o bien C; C si y solo si D; dan lugar a cinco posibles concatenaciones, pero los sujetos tienden a formas la que encaja con los modelos mentales de las premisas: (A si y solo si B) o bien (C si y solo si D).

Para que conozca usted el regusto de lo que los participantes tenían que decir, vamos a considerar un protocolo de un participante que utilizaba la estrategia más frecuente, que era dibujar un único diagrama que representaba las posibilidades compatibles con las premisas. Con el problema de la Tabla 19.1, el participante leyó la primera premisa en voz alta: “hay una canica roja si y solo si hay una canica marrón”, y entonces hizo una inferencia inmediata a partir ella para extraer la conclusión, “Si marrón entonces roja”. Dibujó entonces un diagrama de una posibilidad compatible con esta proposición:

Marrón Roja

Leyó la segunda premisa en voz alta, “Marrón o bien gris”, y añadió un elemento a este diagrama para representar una posibilidad alternativa:

Gris

Marrón Roja

Representó de esa forma las posibilidades independientes en líneas horizontales

independientes. Leyó la tercera premisa: “Hay una canica gris si y solo si hay una canica negra”, e infirió: “Si gris entonces negra.” Añadió un nuevo referente a su diagrama:

Gris Negra

Marrón Roja

El diagrama sostiene la conclusión: si no roja entonces negra. Lo aceptó diciendo: “Sí”. Luego comprobó la inferencia, revisando de nuevo todas las premisas y comparándolas con su diagrama para las dos posibilidades. Algunos participantes que utilizaron esta estrategia dibujaron una línea vertical de arriba a abajo de la página y escribieron los colores de las canicas para las dos posibilidades a uno u otro lado de ella. Otros, como en el caso del que hablamos, distribuyeron las posibilidades en filas horizontales. Un participante dibujó círculos alrededor de los términos de las propias premisas para extraer una de las dos posibilidades. Un signo revelador de la estrategia era que los participantes revisaban las premisas en el orden en que estaban establecidas, e incluían en sus diagramas información a partir de premisas que eran irrelevantes para evaluar la conclusión.

Otra estrategia frecuente era seguir las consecuencias de una única posibilidad, o bien una que fuera una proposición categórica de las premisas, o una que un participante imaginaba en una suposición. En la vida real, a menudo suponemos que una proposición se cumple y luego extraemos las consecuencias de esta suposición. Supongamos que la inflación aumenta a medida que la economía mejora, ¿qué sucede entonces? Utilizamos nuestro conocimiento general para inferir varias consecuencias, p. ej., nuestros ahorros perderán valor. Cuando hacemos una suposición, debe plasmarse en nuestra conclusión. Debemos concluir: si la inflación aumenta, entonces nuestros ahorros perderán valor. Pero, a veces, nuestra conclusión puede volver atrás para mostrar que nuestra suposición debe ser falsa. Supongamos, por ejemplo, que queremos ir a ver una película pero también necesitamos hacer algo de ejercicio. Pensamos:

Si esta noche voy al cine entonces mañana iré a dar un paseo.

Pero estamos muy ocupados, y por tanto no tenemos tiempo de hacer las dos cosas, y por tanto pensamos:

O bien iré al cine esta noche, o bien iré mañana a dar un paseo.

Entonces podemos pensar: Supongamos que voy al cine esta noche. Podemos utilizar nuestros modelos mentales de nuestra premisa condicional para inferir que iremos a dar un paseo mañana. Pero, esta conclusión y nuestros modelos mentales de nuestra premisa disyuntiva dan lugar a la conclusión de que no iremos al cine esta noche. Por tanto, nuestro razonamiento nos ha llevado a la negación de nuestra proposición: si voy al cine esta noche no iré al cine esta noche. La conjunción de ir y no ir al cine da lugar al modelo nulo o vacío, que recordará usted que representa las contradicciones. Pero, si ampliamos los modelos de nuestra conclusión condicional a modelos completos, dan lugar a la negación de nuestra suposición. Hemos “reducido” nuestra suposición de ir al cine a un absurdo, y los lógicos se refieren a este tipo de argumento como reductio ad absurdum. Los psicólogos asumen a veces que para hacerla necesitamos una regla formal de inferencia. Pero, como acabo de demostrar, podemos hacerla utilizando modelos, aunque necesitamos modelos completos para apreciar la fuerza de la contradicción. Por tanto, la inferencia no es trivial, sino que muestra que debemos rechazar nuestra suposición. A su vez, su negación lleva, a partir de nuestra premisa disyuntiva, a la conclusión posterior de que mañana iremos a dar un paseo.

No siempre somos lógicos en el uso de nuestras suposiciones. Dada una conclusión que tenemos que evaluar, como por ejemplo: “Si hay una canica roja entonces hay una canica negra”, el procedimiento correcto en lógica es hacer la suposición de que hay una canica roja y luego tratar de mostrar que se sigue que hay una canica negra. Un error habitual es trabajar en la dirección opuesta: hacer la suposición de que hay una canica negra y mostrar que se sigue que hay una

canica roja. Este procedimiento muestra que si hay una canica negra entonces hay una canica roja, pero la conclusión que hay que probar es la conversa. No solo hacemos suposiciones respecto a las cláusulas de los condicionales. Algunas veces también hacemos suposiciones basadas en disyunciones. Nos diferenciamos unos de otros en nuestra tendencia a hacer suposiciones. Al menos en los experimentos, hay quien las hace raras veces, mientras otros las derrochan, haciéndolas incluso cuando existen premisas categóricas.

La estrategia compuesta es difícil porque exige que tengamos presente una cantidad considerable de información. Da lugar a conclusiones a partir de pares de premisas compuestas, es decir, las que contienen una conectiva proposicional. Una característica de la estrategia es que incluso cuando ni las premisas ni la conclusión utilizan la palabra, “posiblemente”, los participantes a menudo extraen conclusiones sobre una posibilidad. Por ejemplo, un participante hizo esta inferencia compuesta:

O bien hay una canica roja o bien una canica azul.

O bien hay una canica azul o bien una canica gris.

Por tanto, posiblemente hay una canica gris y una canica roja.

La conclusión se refiere solo a una de las dos posibilidades consistentes con las premisas. Reduce la cantidad de información que los participantes tienen que tener presente.

La estrategia en cadena exige convertir todas las premisas, incluso las disyunciones, en condicionales. Es más fácil hacer inferencias a partir de condicionales que a partir de disyunciones, porque los condicionales tienen un modelo mental explícito, mientras que las disyunciones tienen al menos dos modelos mentales explícitos. Por tanto, la teoría de modelos predice que deberíamos construir cadenas de condicionales en lugar de cadenas de

disyunciones, si bien estas últimas se utilizan en ciertos programas de ordenador para la comprobación de teoremas. En realidad, hasta ahora, ningún participante ha convertido nunca un conjunto de premisas en una cadena de disyunciones.

Las cinco estrategias se pueden utilizar con cualquier inferencia proposicional. Pero, dependiendo de los tipos de problemas de un experimento, varía la frecuencia de uso de cada una de ellas. Por ejemplo, los participantes tienden a utilizar la estrategia de cadena para evaluar conclusiones condicionales, mientras que tienden a utilizar la estrategia de concatenación para formular sus propias conclusiones. En todos los experimentos, sin embargo, la estrategia más frecuente era construir un único diagrama de las posibilidades consistentes con las premisas. Los participantes mezclan las estrategias y cambian de una a otra de manera impredecible. Algunas veces cambian en medio de un problema; otras veces, de un problema a otro. No hay secuencias fijas de pasos que todo el mundo siga. No siguen una única estrategia determinista.

* * *

¿Cómo desarrollamos las estrategias? Mi primera asunción (del Capítulo 1) es que el razonamiento no es un proceso determinista que se desenrolla como un mecanismo de relojería. Está gobernado por limitaciones, pero estas raramente dan lugar a un único camino a seguir. Varía de formas que solo pueden expresarse mediante una descripción que no sea determinista. El propio cerebro puede ser o no ser determinista. Podría ser nuestra ignorancia la que hace que parezca que no lo es. Pero la asunción de que el razonamiento no se puede explicar de una forma determinista se corrobora a dos niveles. A nivel superior, desarrollamos diversas estrategias. A nivel inferior, hay variación dentro de nuestro propio razonamiento. Cuando seguimos una estrategia determinada, no seguimos una secuencia fija de pasos tácticos para cada inferencia. Por ejemplo, cambiamos respecto a si leemos una premisa una o dos veces, respecto a qué proposición utilizamos como suposición, si es que hay alguna, y respecto a dibujar o no un diagrama para representar una premisa. Llegamos al laboratorio equipados ya con un conjunto de tácticas inferenciales. Podemos construir un diagrama o un modelo mental, actualizarlo, hacer suposiciones, formular conclusiones, etc. Cuando tratamos de razonar sobre un problema, ponemos a

prueba varias secuencias con ese tipo de pasos tácticos. No estamos gobernados por ningún plan generalizado, preparado con anterioridad, pero tampoco damos palos de ciego. Estamos limitados por la tarea, y por los resultados de nuestras tácticas inferenciales. Si, pongamos por caso, extraemos una conclusión a partir de una suposición y una premisa, esa conclusión nos obliga a considerar si podemos utilizarla para extraer una conclusión posterior. Exploramos diferentes secuencias de tácticas. Estas exploraciones pueden llevarnos no solo a la solución del problema, sino también a una nueva estrategia de razonamiento.

En la jerga de la ciencia cognitiva, componemos una estrategia “desde abajo hacia arriba” [bottom-up] a partir de las exploraciones de los problemas utilizando nuestras tácticas inferenciales. Este método contrasta con el desarrollo “desde arriba hacia abajo” [top-down] de una estrategia que utiliza una planificación de alto nivel. Un método que va de arriba a abajo puede ser posible para los expertos que piensan de forma autoconsciente sobre una rama de la lógica. Sin embargo, nosotros desarrollamos una estrategia que va de los detalles a lo general. Una vez que hemos dominado su uso en una serie de problemas, puede entonces restringir nuestro razonamiento haciendo que vaya de arriba a abajo. Las pruebas corroboran estas hipótesis. Un dato crucial es que aprendemos de nuestros errores. Es decir, podemos utilizar un paso táctico que no lleva a ninguna parte en un problema determinado, pero aprendemos respecto a la táctica, y más tarde nos puede venir muy bien para resolver otro problema.

Una consecuencia de esta explicación es que el conjunto de posibles estrategias viene definido por los diferentes órdenes que se pueden seguir al utilizar las tácticas inferenciales para hacer inferencias. Por tanto, si los psicólogos pueden enumerar todas nuestras tácticas para un tipo determinado de problema, entonces tendrán la base de todas nuestras estrategias factibles para esos problemas. Pero, ¿de dónde surgen las tácticas en sí mismas?

* * *

La teoría de modelos postula que toda táctica inferencial se basa en modelos. Esta consecuencia no excluye el desarrollo de reglas formales, dada la existencia de tácticas que detectan similitudes en cuanto a la forma lógica entre una

proposición y otra. Lo que sí implica, sin embargo, es que nuestro razonamiento depende inicialmente de modelos, y más tarde puede depender de reglas formales. Otra consecuencia es que el tipo de problemas que encontramos debería sesgarnos hacia el desarrollo de estrategias concretas.

Los experimentos han apoyado esta predicción. Por ejemplo, una forma de inculcar el uso de un diagrama único es utilizar premisas que fuerzan a los participantes a pensar en posibilidades alternativas. Un claro candidato es la conectiva "o", y realmente predispone a quien razona a adoptar la estrategia de los diagramas. En cambio, las premisas condicionales dan lugar a un modelo mental explícito, y por tanto tienden a producir la estrategia de cadena o la estrategia de seguir una única posibilidad. Un experimento mostró que la utilización de la estrategia de los diagramas aumentaba cuando los participantes pasaban de problemas basados en condicionales a otros basados en disyunciones, pero no disminuía su utilización cuando cambiaban en sentido contrario. La estrategia de los diagramas es más flexible que las otras estrategias, que dependen de premisas condicionales. Otro experimento mostró que, independientemente de la estrategia, la dificultad de una inferencia depende del número de modelos que exige: había un ocho por ciento de errores con problemas de un modelo, un quince por ciento de errores con problemas de dos modelos y un veinte por ciento de errores con problemas de tres modelos.

* * *

Durante la elección presidencial de 2004 en Estados Unidos, los investigadores informaron de que:

Quienes apoyan a Kerry están mejor informados sobre las opiniones de Kerry de lo que lo están quienes apoyan a Bush sobre las opiniones de Bush.

Esta afirmación expresa una relación entre relaciones. Dice, en realidad, quienes apoyan a Kerry están informados hasta un cierto grado sobre las opiniones de

Kerry, quienes apoyan a Bush están informados hasta un cierto grado de las opiniones de Bush, y el primer grado es mayor que el segundo grado. Este tipo de proposiciones es habitual, p. ej., me gusta más la ópera que el ballet. Las relaciones entre relaciones son un buen "campo de pruebas" para mi hipótesis sobre el desarrollo de estrategias, porque no es probable que, antes de entrar en el laboratorio, los participantes hayan desarrollado una estrategia respecto a ellas. Por tanto, los experimentos han planteado problemas relacionales de este tipo:

Art es más alto que Cal en mayor grado que Bob es más alto que Dan.

Bob es más alto que Dan en mayor grado que Cal es más alto que Dan.

¿Cuál es el orden de los cuatro individuos en términos de altura?

La respuesta en este caso, empezando por el más alto, es: Art, Bob, Cal, Dan. Una estrategia para resolver el problema es darse cuenta de que la segunda premisa establece el orden:

Bob Cal Dan

y luego utilizar la primera premisa para añadir a Art en el lugar adecuado. Si usted duda si Art es más alto que Bob, debería pensar en las consecuencias de suponer que tienen la misma altura. En ese caso es imposible que sea cierta la primera premisa, porque Bob también tendría que ser más alto que Cal en mayor grado que Bob es más alto que Dan, y esto estaría en conflicto con la anterior ordenación de los tres.

Los experimentos mostraron que los participantes desarrollan varias estrategias para enfrentarse con las relaciones entre relaciones. Cambian de una estrategia a otra de una forma impredecible, y determinados tipos de premisas pueden sesgarles hacia el desarrollo de una estrategia en vez de otra. En teoría, no hay límites en la complejidad de las relaciones entre relaciones, Por ejemplo, puede haber una relación entre relaciones entre relaciones:

El margen en que Abe es más alto que Ben en mayor grado que Cal es más alto que Dan es mayor que el margen en que Eve es más alta que Faith en mayor grado que Gerd es más alta que Hope.

Una vez más, el lenguaje puede llevarnos más allá de nuestra capacidad para pensar.

Las pruebas directas sobre estrategias vienen de estudios en los que los participantes piensan en voz alta. A los psicólogos les preocupa el que la tarea de tener que expresar en voz alta nuestros pensamientos pueda afectar a cómo pensamos. Podríamos desarrollar una estrategia que nunca utilizaríamos al razonar sin hablar. Es difícil descartar esta posibilidad, porque cuando razonamos en silencio, los experimentadores no pueden decir qué estrategia estamos utilizando. No obstante, cuando describimos nuestras estrategias con posterioridad, hacemos más o menos la misma narración que al expresar los pensamientos en voz alta, y los patrones en la rapidez y exactitud de nuestras inferencias son más o menos los mismos en ambos regímenes.

* * *

¿Razonan de la misma forma los Kpelle cultivadores de arroz en Liberia que los aldeanos en el remoto Uzbekistán? ¿Razonan los orientales de la misma forma que los europeos? ¿Razonamos nosotros de la misma forma que nuestros ancestros? Los científicos sociales discuten sobre estas cuestiones. Algunos dicen que los seres humanos razonan más o menos igual en todo el mundo y

siempre lo han hecho así; sus oponentes dicen que los seres humanos difieren en cómo razonan dependiendo de su cultura. Algunos sostienen que las normas de racionalidad deberían ser las mismas para todas las culturas; sus oponentes –los llamados “relativistas”– sostienen que las normas deberían diferir y depender de la cultura local. Por tanto, lo que es racional para los Azande de África Central es irracional para Evans- Pritchar, el antropólogo inglés que los estudió, y viceversa. Es difícil resolver las controversias sobre normas, pero la capacidad de educación de los pueblos de todo el mundo sugiere que no existen diferencias profundas en cuanto al razonamiento entre una cultura y otra. El debate se ha formulado a menudo en términos de reglas formales de inferencia. Ninguna regla formal es universal, dicen los relativistas. Pero su argumento presupone que utilizamos reglas formales. Desde el punto de vista de la teoría de modelos, la pregunta a plantear es: ¿se dan cuenta todas las culturas del poder de un contraejemplo? Muéstreme una cultura que no lo haga, y aceptaré que el relativismo es correcto para algunas culturas, pero no para la mía.

El relativismo tiene un aura de autorrefutación (como mi último comentario), porque los relativistas deben admitir que hay una subcultura de racionalistas, que creen que los principios del razonamiento son universales. Supongamos que los relativistas afirman que el racionalismo tiene razón en su subcultura. Se sigue que los principios de razonamiento son universales, porque eso es lo que creen los racionalistas. Y por tanto el relativismo está equivocado. Supongamos que el racionalismo no tiene razón en su subcultura. Se deduce que hay una cultura cuyos principios de razonamiento no son correctos, y por tanto el relativismo está equivocado. De cualquier forma, se deduce que el relativismo está equivocado.

* * *

¿Existen diferencias en los procesos mentales de razonamiento entre una cultura y otra? Hay dos líneas de pruebas psicológicas que resultan sorprendentes. Una muestra que las personas de subculturas sin escritura ni escolarización se muestran reacias a hacer inferencias sobre individuos hipotéticos. Por ejemplo, a un Kpelle cultivador de arroz iletrado de Liberia se le dio, traducido, este problema:

Todos los hombres Kpelle son cultivadores de arroz.

Mr. Smith no es un cultivador de arroz.

¿Es un Kpelle?

Usted o yo responderíamos: no. Pero el granjero respondió: “Si le conozco en persona, puedo responder esa pregunta, pero puesto que no le conozco en persona, no puedo responder a esa pregunta.” Esta respuesta es típica de los individuos pertenecientes a culturas que no tienen escritura. Tienden a hacer inferencias basadas en el conocimiento y la experiencia en lugar de en la sagacidad lógica. Pero incluso los comentarios del granjero muestran que es capaz de hacer deducciones. Está manteniendo: puedo razonar sobre un hombre solo si le conozco; no conozco a Mr Smith; por tanto, no puedo razonar sobre él. Los niños de preescolar o los adultos sin escolarizar de Recife, Brasil, tienen la misma tendencia. Sin embargo, cuando el contexto para la formulación de una inferencia es un planeta distante, donde nadie puede tener ningún conocimiento relevante, la tendencia desaparece y los participantes razonan de forma competente sobre individuos hipotéticos. Asimismo, el efecto de la escolarización o el aprendizaje de la lectura y la escritura es hacer que quienes pertenecen a esas culturas estén mucho más dispuestos a realizar experimentos psicológicos sobre el razonamiento. La diferencia no está tanto en el proceso de razonamiento como en los contenidos sobre los que la gente puede razonar.

La segunda línea de pruebas se debe al psicólogo Richard Nisbett. Él y sus colegas han observado diferencias en el pensamiento entre occidentales y orientales –chinos, japoneses y coreanos–. Son una consecuencia, dice Nisbett, de antiguas diferencias culturales. El pensamiento analítico y el razonamiento deductivo son prerrogativas de los occidentales, mientras que el razonamiento intuitivo holístico que tolera las contradicciones es una prerrogativa de los orientales. Estos investigadores informan de estudios que les llevaron a la conclusión de que los occidentales hacen inferencias formales con más

frecuencia que los orientales, los cuales, como los Kpelle, tienden a razonar a partir de su conocimiento y experiencia.

Un nuevo análisis de estos resultados los ha puesto en duda. Asimismo, algunos estudios sin publicar han sido incapaces de detectar ninguna diferencia en el razonamiento entre personas que hablan chino en Hong Kong y estudiantes occidentales en Princeton. He aquí una prueba de una tarea que resultaba instructiva:

Si un piloto cae de un avión sin paracaídas el piloto muere.

El piloto no murió.

¿Qué pasó?

Los participantes pueden responder o bien con una deducción:

El piloto no cayó de un avión sin paracaídas.

o con una abducción basada en su conocimiento, como por ejemplo:

El avión estaba en el suelo.

La mayoría de las personas da una respuesta abductiva para estos problemas. El sesgo se producía en las respuestas de los chinos y en las de los americanos, y

los dos grupos no se diferenciaron en una prueba estadística. Sin embargo, ambos grupos tenían más tendencia a hacer la deducción si acababan de llevar a cabo un experimento deductivo que no relacionado. En la medida en que existen diferencias culturales en el razonamiento, parecen producirse más en la estrategia que en los mecanismos de razonamiento subyacentes. El propio Nisbett ha comentado que las diferencias son una cuestión de costumbre.

* * *

Cuando tenemos que hacer una secuencia de inferencias desarrollamos nuestro razonamiento a corto plazo. Descubrimos estrategias, no trazando algún gran diseño para el razonamiento, sino probando diferentes pasos tácticos hasta que encontramos una estrategia que alcanza una conclusión. Como predice esta hipótesis, la naturaleza concreta de los problemas nos predispone en favor de un tipo de estrategia u otra. No obstante, independientemente de nuestra estrategia, un invariante es que el número de modelos requeridos para una inferencia afecta a su dificultad. Esta observación sugiere que los mecanismos subyacentes a las tácticas son sensibles al número de posibilidades compatibles con las premisas. Pero eso no implica que solo utilicemos modelos. También nos resulta posible desarrollar una estrategia que se aprovecha de algunas reglas formales. Sin embargo, la invención de la lógica formal es la estrategia final para solo un puñado de individuos excepcionales.

Nuestro razonamiento se organiza en diferentes niveles. En el nivel más alto está el metarrazonamiento –el tipo de razonamiento que nos lleva a evaluar nuestra propia actuación–. Es raro en los protocolos de los experimentos. Nadie señaló, por ejemplo, “Ahora veo cómo puedo resolver estos problemas”, y pasó a describir una estrategia basada en el insight. Pero, en algunas circunstancias, sí que aparecen comentarios sobre el metarrazonamiento, por ejemplo, cuando instruimos a otros sobre cómo razonar. El segundo nivel es estratégico. Una vez que hemos desarrollado una estrategia, su representación mental permite utilizarla en el futuro. Esta representación mental gobierna las secuencias de tácticas que se pueden observar en los experimentos. El tercer nivel es táctico. Consiste en tácticas individuales, como por ejemplo, dibujar un diagrama de una premisa, o añadir información de una premisa a un modelo que ya existe. El

nivel cuarto, e inferior, está formado por los mecanismos que subyacen en las tácticas. No somos conscientes de ellos. Son comparables al “conjunto de instrucciones” de un chip de ordenador, es decir, las operaciones fundamentales que el chip lleva a cabo, y a las que deben traducirse todos los programas de ordenador. La naturaleza de nuestros mecanismos de razonamiento es controvertida, pero las pruebas apoyan la existencia de mecanismos que construyen y manipulan modelos.

Los psicólogos no pueden predecir qué estrategia se desarrollará al tratar de resolver un conjunto de problemas. Una vez que se tiene una estrategia, hay quien parece inflexible en su uso, pero luego cambia de estrategia repentinamente; otros ni siquiera llegan a establecer una opción consistente. Incluso las opciones tácticas son inciertas. Por ejemplo, leemos una premisa en voz alta y pasamos a la siguiente premisa, o podemos leer la premisa de nuevo, y otra vez más, antes de seguir adelante. Los psicólogos deben conformarse con teorías que son de punta a cabo no deterministas.

¿Existe una cultura incapaz de razonar? Lo dudo. Si los antropólogos hubieran descubierto alguna, sería actualmente tan famosa como esas culturas que no tienen más números que “aproximadamente uno”, “aproximadamente dos” y “muchos”. En la medida en que se producen diferencias en el razonamiento parece también que se refieren más a la estrategia que a un profundo arraigo. La popularidad de los Sudokus en ciertas partes de Asia oriental sugiere que su población no tiene problemas para el razonamiento deductivo en contra de la hipótesis de que es una prerrogativa de los occidentales.

¿Por qué entonces se producen grandes diferencias en la visión del mundo entre una cultura y otra? Muchas culturas siguen creyendo en la brujería, la magia y en fuerzas que no tienen una justificación científica. Las culturas difieren en sus creencias, y no existe un antídoto sencillo para las creencias erróneas. Tenemos que ser crédulos para hacer inducciones, y seguramente todos los mecanismos inductivos son falibles. El pensamiento mágico, como sostuve en un capítulo anterior, es general en todas las culturas. El único correctivo que conocemos es la investigación científica, pero es cara y en sí misma depende de conocimiento que todavía tiene que llegar a muchas partes del mundo. Sin embargo, encontramos las semillas de la ciencia y la tecnología en un núcleo de principios racionales comunes, parece ser, a todas las culturas. Se pone de manifiesto en la navegación de los habitantes de Micronesia, en las técnicas de caza de los bosquimanos del Kalahari y en los métodos para formar las puntas de flecha de

piedra de los indígenas americanos. No existe un argumento sólido en favor de las diferencias culturales respecto al proceso de pensar.

CAPÍTULO 20

Cómo podemos mejorar nuestro razonamiento

Las inteligencias sobrehumanas, como he señalado anteriormente en varias ocasiones, tendrían una memoria de trabajo casi ilimitada en la que tener presente información mientras razonaban sobre ella, y pensarían con mucha más rapidez que nosotros. Serían capaces de pensar en muchas posibilidades alternativas, y tanto sobre lo que es verdad como sobre lo que es falso. Nos superarían en razonamiento. Creerían que nuestras capacidades mentales eran defectuosas; nosotros les despreciaríamos por sus extraordinarios poderes. Obtendrían puntuaciones perfectas en todos nuestros test de razonamiento e inteligencia. De hecho, serían mucho más inteligentes que nosotros. Ciertamente, si pudiéramos mejorar nuestro razonamiento, también nosotros actuaríamos mejor en los test y seríamos mejores en las destrezas que predicen los test. Seríamos capaces de resolver problemas que actualmente nos frustran. Seríamos capaces de evitar algunos de los desastres y catástrofes que actualmente nos amenazan. Viviríamos más –los individuos inteligentes lo hacen–. Y podríamos incluso ser más sensatos y vivir en sociedades mejores que las de ahora.

Los escépticos podrían sostener por el contrario que nuestros errores de razonamiento, que he descrito a lo largo de este libro, no importan demasiado. Nos las arreglamos. Sobrevivimos día a día, y cuando nos damos cuenta de nuestros errores, nos recuperamos de ellos. Nuestra existencia ininterrumpida lo prueba. Los errores en el razonamiento, dirían, son un tema de interés exclusivamente académico. Estarían equivocados. Los premios Darwin son un compendio de casos en los que los individuos, en las despiadadas palabras de Wendy Northcutt, han “mejorado la dotación genética eliminándose a sí mismos de la raza humana de una forma asombrosamente estúpida”. Como el hombre que miró qué había dentro de una lata de gasolina con un encendedor, lo hacen sin pensar en una de las posibilidades. Muchas de las catástrofes importantes tienen la misma causa. Consideremos solo un ejemplo que mencioné anteriormente, el hundimiento del barco The Herald of Free Enterprise. Este

ferry para coches se hundió porque sus puertas de proa se habían dejado abiertas al hacerse a la mar. La locura de dirigir un barco de forma que permita este increíble error garrafal se puede atribuir a una causa: los responsables no razonaron bien. Habían establecido un sistema de “informe negativo” en el que la tripulación solo informaba de aquellos sucesos que habían ido mal. No informar de un fallo es un desastre en ciernes, y el antídoto es pedir también informes positivos de que las acciones se han llevado a cabo, p. ej., “las puertas de proa se han cerrado”. El no hacer ese informe habría alertado al capitán del barco del descuido.

¿Nos es posible mejorar nuestro razonamiento? El capítulo anterior mostró que mejoramos a medida que desarrollamos estrategias. Si nos dan un Sudoku extremadamente difícil, no conseguiremos resolverlo. Pero si nos dejan trabajar con otros más fáciles, desarrollaremos –si perseveramos– la estrategia y las tácticas para tener una buen opción de resolverlo. Por tanto, ¿qué podemos hacer para mejorar nuestro razonamiento, hay alguna estrategia general que nos pueda ayudar?

Los métodos que pretenden mejorar el razonamiento son tan abundantes como las dietas, y quizá no más eficaces. Unos son grandiosos como el proyecto de los años 1980 para mejorar la inteligencia de toda la población de Venezuela; otros son tan minúsculos como los libros de autoayuda. Nuestra mayor esperanza –como con la dieta– es preguntar: ¿alguna prueba sólida muestra que el régimen funciona? Si no es así, no estamos obligados a arrojarlo a las llamas, pero deberíamos pensarlo dos veces antes de invertir en él.

* * *

¿Qué pasa con la lógica? Algunos extremistas dicen que no podemos razonar a menos que aprendamos lógica. La moda de los Sudokus muestra que esta afirmación es falsa. Quizá la lógica es una llave para el buen razonamiento. Pero, ¿el estudio de la economía mejoraría nuestra capacidad para manejar el dinero? Algunos economistas se han convertido en expertos inversores –John Maynard Keynes, por ejemplo–, pero otros han sido desastrosos, como los ganadores de un premio Nobel que llevaron a la bancarrota a un fondo de reserva. El entender

teóricamente cómo debemos razonar podría no resultar tan útil en la práctica.

Podemos utilizar la lógica para evaluar argumentos. Thomas Paine, el escritor de panfletos revolucionario, construyó una vez este argumento:

Respecto a lo que se llaman confesiones religiosas, si se deja que cada uno juzgue su propia religión, no existe ninguna religión que esté equivocada, pero si se tiene que juzgar la religión de cada uno de los demás, no existe ninguna religión que sea correcta, y por tanto, todo el mundo tiene razón o todo el mundo está equivocado.

¿Es bueno el argumento? La mayoría de los métodos de lógica utilizan pruebas formales, y la Tabla 20.1 ilustra cómo analizar el argumento por esa vía. Implica tres pasos: resumimos el argumento, recuperamos su forma lógica y establecemos todas las premisas desaparecidas, probamos que la conclusión se sigue de las premisas utilizando reglas formales. Los tres pasos pueden ser difíciles. Mi análisis se inspiró en otro de Geoffrey Keene, pero el suyo es diferente. ¿Qué análisis es correcto? Ambos me parecen verosímiles, y no existe ningún método para decidir sobre sus respectivos méritos. La prueba que aparece en la tabla es fácil de encontrar; otras pruebas –a veces para inferencias simples– pueden ser muy difíciles de descubrir. Un método alternativo para el análisis de argumentos es utilizar modelos completos, y la Tabla 20.2 muestra el resultado que obtiene mi programa para el argumento cuando utiliza modelos completos.

Tabla 20.1. El análisis del argumento informal de Paine utilizando la lógica formal

Paso 1. Resumir el argumento para extraer su forma lógica:

1. Si cada uno juzga su propia religión entonces ninguna religión está equivocada.

2. Si cada uno juzga la religión de los demás entonces ninguna religión es correcta.

Por tanto, todo el mundo tiene razón, o todo el mundo está equivocado.

Paso 2: Simbolizar la forma lógica del argumento, aunque aquí he evitado el uso de notación lógica, y establecer todas las premisas desaparecidas:

1. Si y solo si cada uno juzga su propia religión entonces ninguna religión está equivocada.

2. Si y solo si cada uno juzga la religión de los demás entonces ninguna religión es correcta.

3. O bien cada uno juzga su propia religión o bien cada uno juzga la religión de los demás.

4. Si y solo si ninguna religión está equivocada entonces todo el mundo tiene razón.

5. Si y solo si ninguna religión es correcta entonces todo el mundo está equivocado.

Paso 3: Probar el argumento utilizando reglas formales de inferencia.

1. Supongamos que cada uno juzga su propia religión.

2. Se sigue de la primera premisa que ninguna religión está equivocada.

3. Se sigue de la línea anterior y de la cuarta premisa que todo el mundo tiene razón.

4. Ahora supongamos que cada uno no juzga su propia religión.

5. Se sigue de la tercera premisa que cada uno juzga la religión de los demás.

6. Se sigue de la línea anterior y de la segunda premisa que ninguna religión es correcta.

7. Se sigue de la línea anterior y de la quinta premisa que todo el mundo está equivocado.

8. Se sigue que todo el mundo tiene razón o bien que todo el mundo está equivocado, porque se cumple una suposición o la otra (una inferencia con una forma conocida como un dilema).

Por tanto, la inferencia es válida: si sus premisas son verdaderas, entonces su conclusión es también verdadera. Pero, ¿es verdadera la tercera premisa?

Como muestran ambas tablas, hay un análisis para el que el argumento de Paine es válido. ¿Es verdadera su conclusión? Como siempre, no debemos confundir una conclusión válida con una conclusión verdadera. El que lo sea depende también de si las premisas son verdaderas, y una de las premisas que tuve que añadir al argumento era: O cada uno juzga su propia religión o cada uno juzga la religión de los otros. Esta premisa es falsa: mucha gente no juzga ninguna religión, y algunos pueden juzgar tanto la suya como la de los otros. El mundo es más gris de lo que Paine imaginó.

Un problema al utilizar la lógica es que tenemos que invertir tiempo y esfuerzo en aprenderla. La mayoría de los métodos que se usan para enseñarla ponen el énfasis en las pruebas formales, y por tanto también necesitamos adquirir la capacidad de extraer la forma lógica de los argumentos cotidianos. Un método opuesto para enseñar lógica, iniciado por el difunto Jon Barwise y por John Etchemendy, es poner el énfasis en el contenido en lugar de en la forma. Ellos y sus colegas escribieron un programa de ordenador, “Hyperproof ”, que permite a los estudiantes construir oraciones en un cálculo lógico y evaluar si son verdaderas o falsas en relación con un modelo de bloques distribuido en un tablero de ajedrez. Un estudio psicológico respecto a la eficacia de Hyperproof como forma de enseñar lógica ha mostrado que funciona para algunos estudiantes, pero no para otros, quizá debido a diferencias en su estilo de pensamiento.

Nadie sabe si un curso de lógica mejoraría nuestro razonamiento en la vida real. No debería hacernos daño, pero las pruebas muestran que sus efectos de mejora pueden ser limitados. Cuando los participantes realizaron un curso semestral de lógica proposicional (que trataba de la lógica del “si”, “o” e “y”), su capacidad para hacer frente a tareas de razonamiento basadas en condicionales –la tarea de selección– era mejor que la de un grupo que no siguió el curso de lógica. Dado un condicional, “Si hay una A en un lado de la carta entonces hay un 2 en el otro lado”, los participantes con entrenamiento lógico entendían lo importante que

era una carta que tenía un tres como contraejemplo potencial para el condicional. Sin embargo, cuando la afirmación general se describía en términos de un cuantificador, “Toda carta con una A en una cara tiene un 3 en la otra cara”, su ventaja se evaporaba y no lo hacían mejor que el grupo sin entrenamiento en lógica. En el curso de lógica no se habían abordado los cuantificadores, y los participantes que lo habían seguido eran incapaces de ver que la proposición cuantificada exigía las mismas comprobaciones que la proposición condicional.

¿Cuando el 28 de marzo de 1979 “se desconectó” la turbina, debieron haber traducido los ingenieros a cargo de la central de Three Mile Island su conocimiento a un cálculo lógico y buscado una prueba de la causa de la emergencia? La idea es absurda. La lógica consume demasiado tiempo como para usarla en emergencias. Tenemos que pensar con rapidez. Pero el escepticismo respecto a que la lógica sea una ayuda para razonar no debería confundirse con la hostilidad peligrosísima de algunos expertos. Escriben sobre los peligros de aplicar la lógica a los asuntos humanos, las deducciones estrechas de miras y las conclusiones poco razonables. Gran parte de su retórica se apoya solo en una confusión entre verdad y validez. Pura tontería, incluso si una deducción es válida.

Tabla 20.2. El análisis del argumento informal de Paine utilizando modelos completos

Las premisas y la conclusión, donde “sis” representa “si y solo si” y “ob” representa “o bien” (una disyunción excluyente), son las siguientes:

Cada uno juzga su propia religión ob cada uno juzga la religión de los demás.

Sis cada uno juzga su propia religión entonces ninguna religión está equivocada.

Sis cada uno juzga la religión de los demás entonces ninguna religión es correcta.

Sis ninguna religión está equivocada entonces todo el mundo tiene razón.

Sis ninguna religión es correcta entonces todo el mundo está equivocado.

Todo el mundo tiene razón ob todo el mundo está equivocado (la conclusión).

El resultado del programa para cada premisa y conclusión (en forma abreviada en la que "¬" representa "no"):

Juzga la propia ob juzga la de los demás, da lugar a los modelos completos:

juzga-propia ¬ juzga-demás

¬juzga-propia juzga-demás

Sis juzga propia entonces nadie equivocado, actualiza los modelos a:

juzga-propia ¬ juzga-demás nadie-equivocado

¬juzga-propia juzga-demás ¬ nadie-equivocado

Sis juzga demás entonces nadie tiene razón, actualiza los modelos a:

juzga-propia ¬ juzga-demás nadie-equivocado ¬ nadie-tiene-razón

¬juzga-propia juzga-demás ¬ nadie-equivocado nadie-tiene-razón

Sis nadie equivocado entonces todos tienen razón, actualiza los modelos a:

juzga-propia ¬ juzga-demás nadie-equivocado ¬ nadie-tiene-razón todos-tienen-razón

¬juzga-propia juzga-demás ¬ nadie-equivocado nadie-tiene-razón ¬ todos-tienen-razón

Sis nadie tiene razón entonces todos equivocados, actualiza los modelos a:

juzga-propia ¬juzga-demás nadie-equivocado ¬nadie-tiene-razón todos-tienen-razón ¬ todos- equivocados

¬ juzga-propia juzga-demás ¬ nadie-equivocado nadie-tiene-razón ¬ todos-tienen-razón todos- equivocados

Todos tienen razón ob todos equivocados, da lugar al resultado:

Tu argumento implica VÁLIDAMENTE el aserto: Todos tienen razón ob todos equivocados

* * *

Muchos de los que enseñan a razonar a los estudiantes han abandonado la lógica como herramienta útil. Una de sus preocupaciones legítimas es la absoluta dificultad del análisis lógico. Otra es la distancia entre la lógica y las inferencias cotidianas. El filósofo Stephen Toulmin sostenía que la lógica es inadecuada para el análisis de los temas reales. Las premisas, decía, tienen papeles diferentes en los argumentos, y la lógica no consigue distinguirlos. Muchos pedagogos compartían este escepticismo, y formaron una asociación en favor de la lógica informal y el pensamiento crítico (AILACT). Sus miembros defienden todo tipo de métodos de análisis –la identificación de falacias, la utilización de diagramas especiales o programas de ordenador, y muchos otros más–. Estos diversos métodos no son mutuamente excluyentes, pero han sido mutuamente explosivos. Quienes proponen un sistema consideran que los otros sistemas son impracticables. Y nadie parece haber demostrado mejoras importantes en el razonamiento como resultado de ninguno de ellos.

* * *

Muchos autores han sostenido que los diagramas nos pueden ayudar a razonar. La pregunta sobre cómo podrían hacerlo tiene una larga historia. Ya he descrito los círculos de Euler y los diagramas de Venn para la representación de silogismos. Lewis Carroll desarrolló un juego de lógica para el mismo dominio. Sin embargo, el sistema de diagramas más potente sin lugar a dudas se debe a

Peirce. Lo diseñó, no con propósitos pedagógicos, sino para mostrar los pasos de nuestro pensamiento. Una vez más, nadie sabe si nuestro razonamiento se beneficiaría del aprendizaje de cualquiera de estos sistemas.

En un estudio psicológico pionero, el difunto Herb Simon y sus colegas diferenciaron diversas formas de que los diagramas sean útiles. Pueden facilitar el encontrar información relevante, porque los podemos recorrer rápidamente con la vista. Pueden facilitar la identificación de ejemplos de un concepto, porque podemos reconocer un icono más rápidamente que una descripción. Pero, según Simon, no deberían producir ninguna diferencia en el razonamiento. Este debería funcionar de manera muy similar tanto si las premisas son oraciones como si son diagramas. Barwise y Etchemendy también sostuvieron que los diagramas pueden presentar una multitud de detalles contenidos en las conjunciones, p. ej., el vehículo tiene seis ruedas y una torreta montada en la parte atrás y es descapotable y tiene un gran parachoques. Pero, decían, los diagramas resultan pobres para representar premisas disyuntivas o negativas (basadas en “o”), las cuales se presentan mejor mediante oraciones. Su programa Hyperproof utiliza esta división del trabajo entre diagramas y oraciones.

La teoría de modelos hace una predicción diferente. La mera utilización de un diagrama no es una garantía de que nuestro razonamiento mejorará. Pero los diagramas que nos ayudan a imaginar diferentes posibilidades deberían mejorar nuestro razonamiento, incluso a partir de negaciones y disyunciones. Puedo ilustrar este punto con inferencias complicadas a partir de pares de premisas disyuntivas, como por ejemplo:

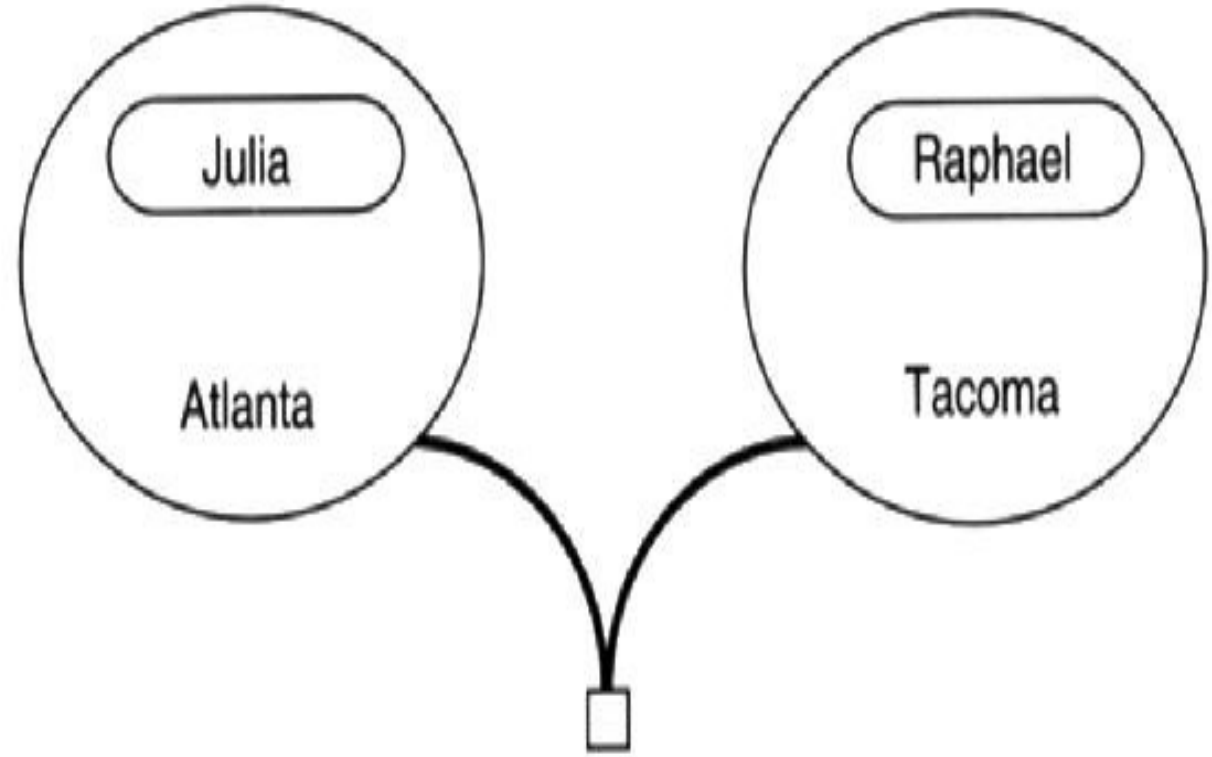
Julia está en Atlanta o Rafael está en Tacoma, o ambas cosas.

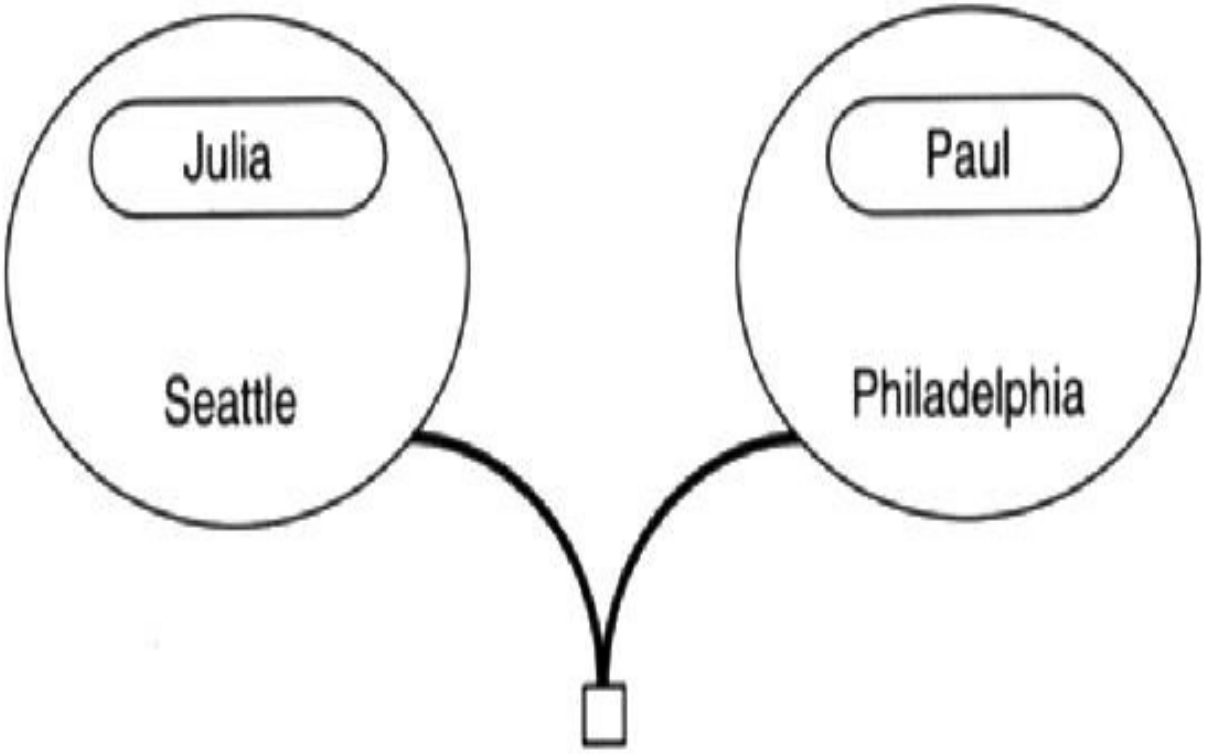
Julia está en Seattle o Paul está en Filadelfia, o ambas cosas.

¿Qué se sigue, si se sigue algo?

Dado que Julia no puede estar en ambos sitios, Atlanta y Seattle, se deduce que Rafael está en Tacoma o Paul está en Filadelfia, o ambas cosas. Pero la inferencia es difícil, porque tenemos que tener presentes demasiadas posibilidades. ¿Podría hacerlo más fácil un diagrama?

Un primer estudio comparó inferencias partiendo de premisas verbales con otras partiendo de diagramas. La Figura 20.1 muestra uno de los diagramas del experimento, que representa las premisas anteriores. Las elipses representan la localización de una persona en una ciudad, y las dos líneas que conectan las elipses se cruzan en un cuadrado. El cuadrado representa una disyunción incluyente, que permite ambas posibilidades. Un símbolo diferente, que no se muestra en la figura, representa una disyunción excluyente, que no permite que se produzcan ambas posibilidades. El experimento confirmó que las disyunciones excluyentes, que dan lugar a dos posibilidades, son más fáciles que las disyunciones incluyentes, que dan lugar a tres posibilidades. Sin embargo, los diagramas no mejoraron el razonamiento. Si examina usted de nuevo la Figura 20.1, puede ver que los diagramas no consiguen hacer explícitas las diferentes posibilidades compatibles con las premisas. Y no consiguen hacer explícita la fuerza de una negación implícita: en el diagrama no hay nada que represente que si Julia está en Atlanta, entonces no está en Seattle, y viceversa.



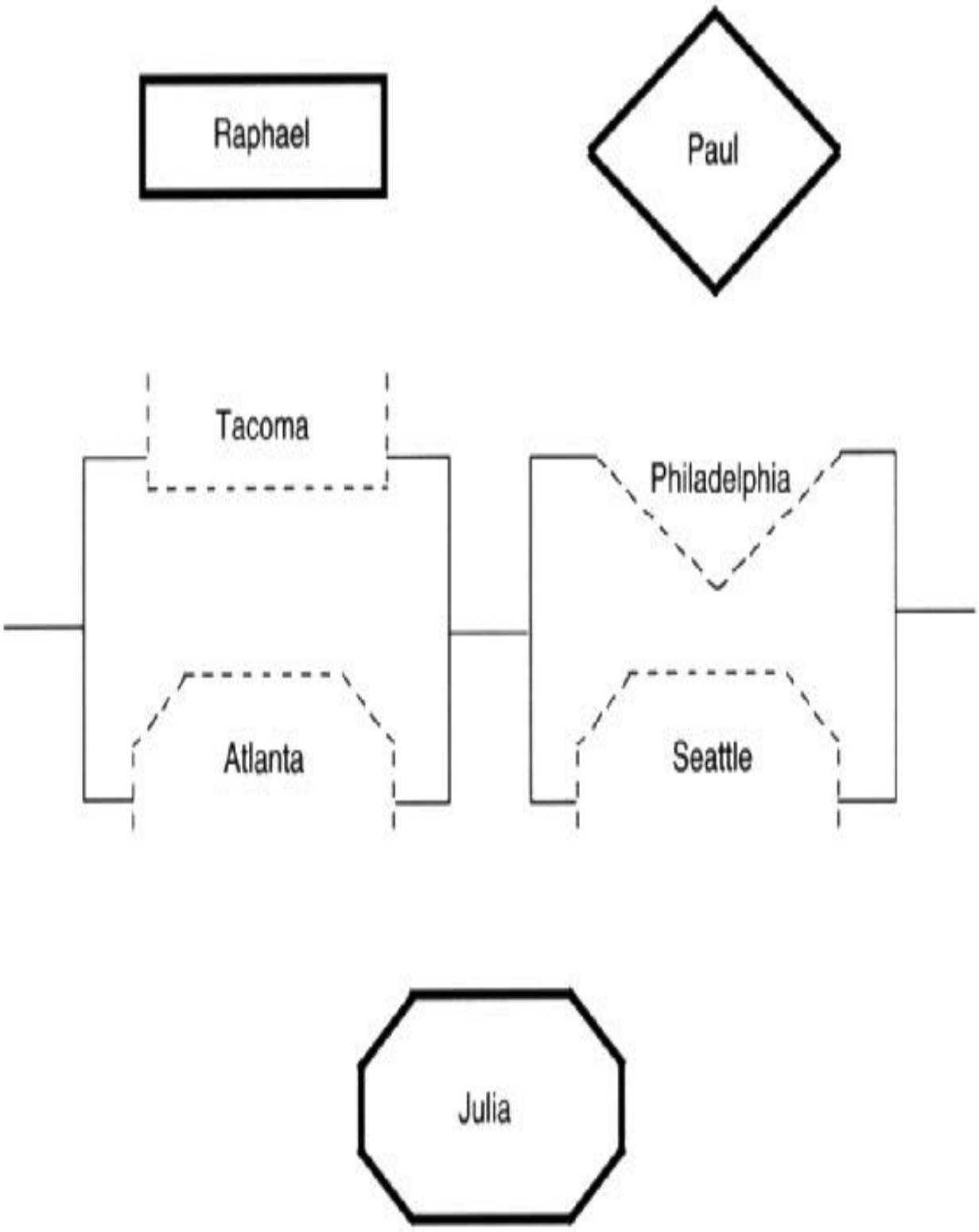


What follows?

Fig. 20.1. Un diagrama utilizado para presentar las premisas en el estudio preliminar

Un segundo experimento utilizó diagramas muy diferentes con el fin de corregir estas deficiencias. Un grupo de participantes razonó a partir de los nuevos diagramas y otro grupo de participantes razonó a partir de premisas verbales. A los participantes se les dijo que las premisas tenían que ser verdaderas para que se produjera un determinado suceso. La Figura 20.2 muestra el diagrama que presenta el mismo problema de antes. Los participantes sabían que tenían que completar una ruta desde un lado de la figura al otro insertando las formas que representan a las personas en los espacios que representan a los lugares. Como muestra la figura, la forma que representa a una persona solo puede encajar en el espacio correspondiente a un lugar que tiene Rafael esa misma forma. Por tanto, la forma que representa a Rafael puede estar en Tacoma o no, la forma que representa a Paul puede estar en Filadelfia o no y la forma que representa a Julia puede estar en Atlanta o en Seattle, o en ninguno de los dos.

Los nuevos diagramas aumentaron la exactitud de manera sorprendente. A partir de ellos, los participantes extrajeron un setenta y cuatro por ciento de conclusiones correctas frente al cuarenta y seis por ciento de conclusiones correctas para las premisas verbales equivalentes. También fueron mucho más rápidos al responder a los diagramas (una media de 99 segundos) que al responder a las premisas verbales (una media de 135 segundos). En ambas situaciones, las disyunciones excluyentes (dos posibilidades) fueron más fáciles que las disyunciones incluyentes (tres posibilidades) y los problemas afirmativos fueron más fáciles que los problemas que contenían una negación implícita. Estos patrones se ponían de manifiesto tanto en los porcentajes de respuestas correctas como en los tiempos de respuesta. Las conclusiones erróneas tendían a ser consistentes con alguna de las posibilidades compatibles con las premisas. Esta tendencia era mayor para los problemas con diagramas que para los problemas verbales. Una vez más, los diagramas hicieron más fácil el captar al menos algunas de las posibilidades.



The event is occurring.
What follows?

Fig. 20.2. Un diagrama icónico utilizado para presentar las premisas en el experimento principal.

¿Por qué estos diagramas son tan útiles? La Figura 20.2 es icónica, y por tanto hace que nos sea más fácil imaginar las posibles rutas desde un lado del diagrama al otro. Los diagramas son estáticos, pero podemos imaginarnos moviendo esta pieza o aquella al espacio adecuado. Muestran situaciones negativas, pero podemos imaginarnos deslizando una pieza al espacio adecuado, para representar la correspondiente proposición afirmativa. Resumiendo, los diagramas nos permiten visualizar las posibilidades inherentes a las premisas: traducen la disyunción y la negación a analogías espaciales.

Cuando razonamos a partir de una disyunción, tenemos que captar su significado y utilizarlo para construir modelos de las posibilidades alternativas. Tenemos que actualizar estos modelos con las alternativas de las premisas posteriores. Este proceso es difícil, y podemos perder la pista de dónde estamos. Con un diagrama, sin embargo, formamos una representación icónica, y el ojo de nuestra mente puede visualizar el desplazamiento de las piezas a sus posiciones. Llevamos a cabo una transformación visual, pero es equivalente a imaginarnos una posibilidad. Visualizamos el desplazamiento de la loseta que designa a Julia al espacio que representa Atlanta, y sabemos que el resultado representa la proposición de que Julia está en Atlanta. Podemos describir esta posibilidad en nuestra conclusión. El proceso es más rápido y más exacto que uno basado en premisas verbales. En resumen, no todos los tipos de diagramas mejoran el razonamiento, pero los que nos facilitan el imaginar posibilidades resultan útiles.

* * *

Para mejorar nuestro razonamiento, necesitamos una memoria de trabajo más capaz y quizá procesos mentales más rápidos. Necesitamos mejorar nuestra comprensión de las premisas, nuestra habilidad para pensar en todas las posibilidades compatibles con ellas, nuestra potencia para formular conclusiones

que capturen lo que es común a esas posibilidades y nuestra destreza para encontrar contraejemplos para las conclusiones. Sobre todo, necesitamos más imaginación para prever lo que es posible: Julia puede no estar en Atlanta; las puertas de proa pueden no haber sido cerradas.

Podemos aprender formas de codificar la información en la memoria de trabajo más eficaces, pero el procedimiento exige mucha práctica y solo funciona para conjuntos limitados de materiales. Parece ser que no podemos aumentar la capacidad de procesamiento de la memoria de trabajo en sí misma o la velocidad de las respuestas neurales. Parece ser una consecuencia de la estructura de nuestro cerebro. Por tanto, necesitamos una panacea. Necesitamos un método práctico que se tarde pocos minutos en aprender y sea fácil de utilizar, que enraíce con lo que nos resulta natural, que no requiera aparatos –ni siquiera lápiz y papel, que mejore nuestro razonamiento, que tenga efectos duraderos, y que funcione para cualquier dominio–. Describiré el mejor método que conozco que de alguna manera llega a cumplir estos objetivos.

Al desarrollar estrategias para razonar, hay quien utiliza espontáneamente un único diagrama que representa todas las posibilidades compatibles con las premisas. Como describía el capítulo anterior, hay varias formas de dibujar estos diagramas. Hay quien dibuja una línea vertical en medio de la hoja de papel para separar las posibilidades. Quien dibuja una línea horizontal. Quien dibuja círculos alrededor de aquellos ítems de las premisas que aparecen en una de las posibilidades. Esta estrategia de razonamiento se basa en un diagrama diseñado para seguir el rastro de las posibilidades. Y los datos mostraron que era una estrategia más flexible que otras. Se puede aprender un método de razonamiento basado en esta estrategia en un par de minutos. El llamado “método de modelos” se compone de una sola instrucción: Intente construir todas las posibilidades consistentes con la información dada. Este precepto se puede poner en práctica de una forma simple, que ilustraré con un ejemplo.

Imaginemos que ciertos profesores tienen que impartir un curso, y estas premisas determinan quién impartirá clase en el curso:

Imparte clase el Dr. Comer o imparte clase el Dr. Sugarman, o imparten clase ambos.

Si el Dr. Sugarman no imparte clase entonces imparte clase el Dr. Glucksberg.

El Dr. Hoebel imparte clase si y solo si imparte clase el Dr. Sugarman.

El Dr. Prentice imparte clase.

El Dr. Glucksberg tiene otro grupo y por tanto no puede impartir clase.

¿Quién imparte clase?

Hay tres posibilidades consistentes con la primera premisa, y podemos representarlas en tres columnas separadas por líneas verticales utilizando la primera letra de cada nombre:

C C

S S

La segunda premisa exige que se añada al Dr. Glucksberg al menos en cualquier posibilidad en la que no aparezca el Dr. Sugarman:

C C

S S

G

La tercera premisa exige que se añada al Dr. Hoebel en cualquier posibilidad en la que aparezca el Dr. Sugarman:

C C

S S

G

H H

La cuarta premisa exige que se añada al Dr. Prentice en todas las posibilidades:

C C

S S

G

H H

P P P

La premisa final elimina cualquier posibilidad en la que aparezca el Dr. Glucksber, y por tanto se la puede borrar, y por ello he tachado la columna de la izquierda:

C C

S S

G

H H

P P P

Este diagrama da lugar a una conclusión válida:

Los doctores Sugarman, Hoebel, y Prentice deben impartir clase, y el Dr. Comer puede impartir clase.

Aquí acaba la descripción de mi método.

Para ver si el método era efectivo, un experimento valoró a los participantes con dos bloques de problemas de razonamiento en los que evaluaban conclusiones dadas. A los participantes de un grupo se les dejó que utilizaran sus propios recursos en ambos bloques; a los de otro grupo se les enseñó (casi con las mismas palabras que acabo de utilizar) el método de modelos después del primer bloque de problemas, y a los de otro grupo se les enseñó el método antes del primer bloque de problemas. Los efectos fueron sorprendentes. Sin la ayuda del método de modelos, los participantes hicieron correctamente alrededor de dos tercios de las pruebas, y emplearon una media de veinticuatro segundos en evaluar cada inferencia. Sin embargo, con ayuda del método hicieron correctamente el noventa y cinco por ciento de las inferencias y emplearon una media de quince segundos en evaluar cada inferencia.

Un experimento posterior mostró una mejora comparable cuando no se permitía a los participantes utilizar lápiz y papel. Se les enseñó el método como antes, pero se les dijo que imaginaran las diferentes posibilidades. Los problemas eran un poco más fáciles que los del primer experimento, y el grupo al que se le dejó que utilizara sus propios recursos realizó correctamente el ochenta por ciento de las inferencias. El grupo al que se le enseñó el método de modelos lo hizo mejor: realizaron correctamente el noventa y tres por ciento de las inferencias. Algunos problemas se basaban parcialmente en premisas bicondicionales (“si y solo si”) y otros se basaban parcialmente en disyunciones excluyentes (“o bien”). Como predice la teoría de modelos, las inferencias a partir de bicondicionales eran más rápidas que las que se hacían a partir de disyunciones. Sin embargo, el efecto del aprendizaje del método de modelos aceleraba más las respuestas a las premisas disyuntivas. Es difícil tener presentes dos posibilidades alternativas, y por tanto el que se nos haya enseñado a intentar seguir el rastro de las posibilidades supone una ayuda.

Consideremos este problema:

Allan pertenece a la comisión o bien Bobby pertenece a ella.

Bobby pertenece a la comisión o bien Nigel pertenece a ella.

Nigel pertenece a la comisión si y solo si Wally pertenece a ella.

¿Se sigue que Allan pertenece a la comisión si y solo si Wally pertenece a ella?

Un método alternativo para evaluar conclusiones se basa en hacer suposiciones. Supongamos, por ejemplo, que Allan pertenece a la comisión. De la primera premisa se sigue que Bobby no pertenece a la comisión. De esta conclusión y de la segunda premisa se sigue que Nigel pertenece a la comisión. De esta conclusión y de la tercera premisa se sigue que Wally pertenece a la comisión. Por tanto, hemos establecido que si Allan pertenece a la comisión entonces Wally también pertenece. Pero la conclusión también afirma que Allan pertenece a la comisión solo si pertenece Wally. Para comprobar esta afirmación, necesitamos hacer una suposición más. Una forma de actuar es suponer que Wally pertenece a la comisión. De esta suposición y de la tercera premisa se sigue que Nigel pertenece a la comisión. De esta conclusión y de la segunda premisa se sigue que Bobby no pertenece a la comisión. De esta conclusión y de la primera premisa se sigue que Allan pertenece a la comisión. Por tanto, ahora sabemos que si Wally pertenece a la comisión entonces Allan también pertenece. Hemos mostrado que la conclusión se sigue de las premisas.

Un experimento examinó la eficacia de enseñar a los participantes a utilizar suposiciones, dado que algunas personas tienen una tendencia natural a utilizar esa estrategia. El experimento mostró que aquellos a los que se les enseñaba el método de modelos eran los más exactos (más del noventa por ciento correcto).

Aquellos a los que se les enseñaba a utilizar suposiciones eran los siguientes en exactitud (aproximadamente el ochenta por ciento correcto). Y aquellos a los que se dejaba que utilizaran sus propios recursos eran los menos exactos (aproximadamente el sesenta y cinco por ciento correcto).

Los experimentos mencionados hasta ahora exigían que los participantes evaluaran conclusiones dadas, pero un experimento posterior mostró que el método de modelos también era útil cuando los participantes tenían que formular sus propias conclusiones. Todas las inferencias se basaban en dos premisas, una contenía una conectiva proposicional y otra una proposición categórica sencilla. Aquellos participantes a los que se les enseñó el método de modelos extrajeron muchas más conclusiones válidas (ochenta por ciento) que aquellos a los que se les dejó que utilizaran sus propios recursos (cincuenta y dos por ciento).

La conclusión a partir de estos estudios es que el método de modelos, aunque solo se tarda unos minutos en enseñarlo, tiene fuertes efectos tanto sobre la exactitud como sobre la rapidez de razonamiento. Nos ayuda a tener presentes las posibilidades alternativas compatibles con las premisas. No es solo un efecto general de la enseñanza, porque conduce a una mejora del razonamiento mayor que la que se produce al enseñar a utilizar suposiciones. No sé si produce una mejora de larga duración, porque los estudios no se orientaron hacia esa cuestión, pero el método es eficiente y efectivo a corto plazo. No es infalible, pero explota la tendencia que todos tenemos a pensar en posibilidades.

* * *

A pesar de la existencia de muchos sistemas para mejorar el razonamiento, pocos estudios empíricos han examinado su valía. Los psicólogos ni siquiera saben si el aprender lógica da lugar a que razonemos mejor en la vida real. El estudio que he descrito mostró que no se conseguía transferir ningún efecto beneficioso en relación con inferencias de la misma forma lógica pero diferente formulación. Esta falta de pruebas ha llevado a quienes proponen “la lógica informal y el pensamiento crítico” a abogar por otras formas de considerar los argumentos. Algunos de estos sistemas utilizan diagramas para representar argumentos, lo mismo que hacen muchos sistemas de análisis lógico. Los resultados de los que

he informado sugieren que solo serían útiles ciertos tipos de diagramas. Para ser eficaces, necesitan ayudarnos a imaginar las posibilidades compatibles con las premisas. Otros diagramas a los que les falta esta propiedad icónica no tienen efectos beneficiosos. El “método de modelos” se basa en enseñarnos a pensar sobre las posibilidades compatibles con las premisas. Tiene las virtudes de la simplicidad y la eficacia. El método se aprende rápidamente –se construye sobre nuestra propensión natural a pensar en posibilidades– y mejora tanto nuestra exactitud como nuestra rapidez de razonamiento. La mayoría deberíamos sacar provecho de este método.

Parte VII

Conocimiento, creencias y problemas

Todo conocimiento que se contradice a sí mismo es Ficción Poética

Auden

CAPÍTULO 21

Los misterios del si

“Hasta los cuervos del tejado graznan acerca de la naturaleza de los condicionales”, escribió Calímaco, el bibliotecario de Alejandría en el siglo tercero antes de Cristo. Por tanto, “si” ha estado bajo discusión durante mucho tiempo. Es una palabra problemática. He descrito los principales principios de la teoría de modelos para el razonamiento –desde los mecanismos que construyen modelos hasta el desarrollo de estrategias y mi intención en esta parte del libro es mostrar cómo puede utilizarse la teoría para resolver diversos misterios–. Empiezo por los misterios del si.

Supongamos que usted es un ingeniero que está llevando a cabo una prueba de un sistema complejo y sabe que para que la prueba pueda continuar el sistema no debe haber alcanzado el nivel crítico. Observa entonces que ha alcanzado el nivel crítico. ¿Qué debería hacer? Parece obvio que usted debería detener la prueba. Los ingenieros a cargo de Chernobyl estaban en esa situación, pero continuaron con la prueba. El por qué continuaron es desconcertante, porque la prueba no solo era peligrosa sino inútil. Condujo al desastre.

Durante varios años, he planteado a estudiantes de ingeniería premisas de tipo similar a las que afrontaron los operadores del reactor nuclear de Chernobyl, pero con un contenido más abstracto:

Si en el tablero hay un triángulo entonces en el tablero hay un círculo.

En el tablero no hay un círculo.

¿Qué se sigue, si se sigue algo?

Podría usted pensar en un principio que de las premisas no se sigue nada. Cada año, al menos la mitad de los ingenieros también piensa eso. Quizá los ingenieros de Chernobyl estaban en un estado mental similar. En realidad, hay una inferencia válida. Se sigue que en el tablero no hay un triángulo. La inferencia es bastante difícil para los estudiantes, pero no porque sean reacios a hacer inferencias a partir de condicionales. Con premisas de este tipo:

Si en el tablero hay un triángulo entonces en el tablero hay un círculo.

En el tablero hay un triángulo.

casi todos extraen la conclusión válida de que en el tablero hay un círculo. La diferencia de dificultad entre los dos tipos de inferencia –esta y la inferencia de Chernobyl– es uno de los efectos más sólidos en los experimentos sobre razonamiento. Una infinidad de experimentos la han corroborado.

El razonamiento a partir de si es un lugar común, debido a la gran parte de nuestro conocimiento que tiene forma condicional. Si nos pillan circulando muy rápido pagamos una multa. Si nos operan necesitamos tiempo para recuperarnos. Si tenemos dinero en la cuenta bancaria podemos cobrar un cheque. Por tanto, ¿por qué algunas veces los ingenieros se equivocan en el laboratorio y en la sala de control? Este es el primer misterio del si.

Hay un misterio más profundo. Como Calímaco insinuaba, las proposiciones condicionales son desconcertantes. Han seguido desconcertando a filósofos, lingüistas y psicólogos. Ninguna otra palabrita ha provocado tantos libros. La raíz del problema es fácil de definir. ¿Qué significan los condicionales? Puede usted replicar: es obvio lo que significan –incluso los niños pequeños lo saben–. Entienden muy bien cuando su padre les dice: “Si eres bueno entonces puedes tomar palomitas.” Un niño de tres años respondió a esta promesa de su padre con

el comentario: “Soy bueno”, invitándole a sacar la conclusión obvia. Por tanto, soy comprensivo con su escepticismo. Los niños entienden los condicionales. Quienes no lo hacen son los filósofos, los lingüistas y los psicólogos. Ciertamente, todo el mundo entiende los condicionales en el flujo de la conversación, pero nadie parece entenderlos al reflexionar sobre ellos con tranquilidad analítica.

Los condicionales son inconstantes. A veces significan una cosa; a veces significan otra. Como ejemplo, considere la diferencia entre si fuera el dueño del coche entonces lo conduciría, y si condujera el coche entonces lo estrellaría. Parecen similares pero son bastante diferentes. El primer condicional es consistente con tres posibilidades:

Era el dueño del coche.	Lo conducía.
No era el dueño del coche.	Lo conducía.
No era el dueño del coche.	No lo conducía.

Pero el segundo condicional solo es consistente con dos posibilidades:

Conducía el coche.	Lo estrelló.
No conducía el coche.	No lo estrelló.

La diferencia es que no puede haber estrellado el coche si no lo conducía, mientras que podría haberlo conducido sin ser el dueño. Pocas veces nos damos

cuenta de estas diferencias porque nuestros modelos mentales de los condicionales solo hacen explícita una posibilidad, p. ej.:

Era el dueño del coche. Conducía el coche.

...

Representamos las otras posibilidades en un modelo que no tiene contenido explícito (indicado por la elipsis), y por tanto no nos damos cuenta de que es diferente de un condicional al otro.

Hay muchas otras interpretaciones para los condicionales, y su comportamiento camaleónico ha llevado a algunos teóricos a proponer que la palabra si es ambigua, aunque las ambigüedades no son del tipo de las que reconocen los autores de diccionarios. Otros han tratado a los condicionales como Procrusto trató a sus víctimas. Eliminan los trozos que no encajan en sus teorías. Casi todos han considerado únicos los condicionales, y por eso es por lo que hay libros sobre el si, pero no sobre el o.

La resolución del misterio, creo, es que el si tiene diversas interpretaciones porque en su interpretación interactúan varios componentes sencillos. Estos componentes aparecen también en diferentes combinaciones en la interpretación de otras conectivas. Una vez que he concretado de esta forma el significado de si, surge una descripción del razonamiento más sencilla para explicar por qué la inferencia de Chernobyl es difícil. Describiré estos componentes, esbozando explicaciones alternativas para ponerlos de relieve y resumiendo algunos datos que ayudan a clarificar la teoría.

* * *

El primer componente que influye en los condicionales es la gramática. Si se diferencia de y y o. Estos últimos pueden conectar cláusulas principales, p. ej.:

ella se movió y la radiografía está borrosa. Pero los condicionales tienen una cláusula si subordinada y una cláusula entonces principal. Una consecuencia es que expresiones que parecen aplicarse al condicional como un todo de hecho se aplican solo a su cláusula principal. Oraciones como por ejemplo:

Es probable que si ella se movió entonces la radiografía esté borrosa.

Es obligatorio que si él ganó un sueldo entonces pague impuestos.

tienen por tanto el mismo significado que:

Si ella se movió es probable que la radiografía esté borrosa.

Si el ganó un sueldo entonces es obligatorio que pague impuestos.

Por el contrario, la conjunción es probable que ella se moviera y la radiografía esté borrosa, no tiene el mismo significado que ella se movió y es probable que la radiografía esté borrosa. La primera oración solo afirma que es probable que ella se moviera, pero la segunda oración afirma que ella se movió. Esta migración de las calificaciones de las oraciones como un todo a sus cláusulas principales se producen con otros tipos de oraciones que contienen cláusulas subordinadas, p. ej., es obligatorio que después de emigrar te registres en la policía significa lo mismo que después de emigrar es obligatorio que te registres en la policía. Veremos más adelante que el hecho de la migración puede haber inducido a error a los teóricos respecto al significado de los condicionales.

La cláusula si en un condicional describe una posibilidad, casi siempre de una forma incompleta. La cláusula entonces se interpreta como si fuera una cláusula

principal aislada en el contexto de esta posibilidad. Se sigue que la cláusula si debe ser declarativa con el fin de describir una posibilidad. Pero la cláusula entonces no tiene ese tipo de restricción: puede ser interrogativa y plantear una pregunta, o puede ser imperativa y hacer una petición, p. ej., Si ella se movió entonces haz otra radiografía. Ese tipo de restricción no se produce en las conjunciones y las disyunciones.

Los verbos auxiliares que aparecen en los condicionales también son especiales. El subjuntivo y el condicional perfecto tienen un significado especial en los condicionales, p. ej.: Si hubiera llovido ayer entonces habría habido charcos en la carretera. Según una interpretación de este ejemplo el que haya llovido o no es una cuestión abierta. Según otra interpretación no llovió, y la cláusula entonces describe lo que habría sucedido en la posibilidad "contrafactual" en la que sí llovió.

* * *

La interpretación central de los condicionales es aquella a la que me referí en un capítulo anterior como "lógica". El condicional anterior, si fuera el dueño del coche entonces lo conduciría, tiene esa interpretación. Se produce en diversos condicionales, y en particular en aquellos cuya interpretación no depende del contexto y no evocan ningún tipo de relación entre la cláusula si y la cláusula entonces, más allá de su aparición simultánea en el mismo condicional. La interpretación lógica se produce por tanto para la proposición:

Si en el tablero hay un círculo entonces en el tablero hay un triángulo.

La posibilidad más destacada compatible con esta proposición es la que corresponde a este modelo mental explícito:

En el tablero hay un círculo y un triángulo.

Pero el condicional permite que no haya un círculo en el tablero, y por tanto con pensar un poco la mayoría de nosotros podemos pensar en otras dos posibilidades:

No hay un círculo y hay un triángulo.

No hay un círculo y no hay un triángulo.

Estas tres posibilidades corresponden a los modelos completos de la interpretación lógica de los condicionales, y recordará usted que tenemos tendencia a enumerarlos cuando se nos pregunta por lo que es posible partiendo del condicional.

Hacemos un tipo de juicio diferente cuando evaluamos la verdad o falsedad de ese mismo condicional (Si en el tablero hay un círculo entonces en el tablero hay un triángulo). Si vemos un círculo y un triángulo en el tablero, juzgamos que el condicional es cierto. Si en lugar de eso vemos un círculo sin un triángulo, juzgamos que el condicional es falso. Pero supongamos que vemos solo un cuadrado, o un cuadrado y un triángulo, ¿qué decimos entonces? La mayoría decimos que estos casos son irrelevantes para la verdad o falsedad del condicional, incluso aunque los enumeramos como posibles cuando se nos dice que el condicional es cierto. Ese es otro misterio sobre el si: lo que pensamos que es posible dado un condicional no coincide totalmente con lo que pensamos que hace que un condicional sea verdadero.

La teoría de modelos explica la diferencia. Para enumerar las posibilidades, todo lo que tenemos que hacer es entender un condicional y pensar en lo que es posible. Enumeramos los modelos completos del condicional. Para evaluar si un condicional es verdadero o falso, sin embargo, tenemos que compararlo con una situación que vemos, y también tenemos que entender los términos

“metalingüísticos” verdadero y falso. En este caso es más probable que nos apoyemos en modelos mentales que en modelos completos. Los modelos mentales del condicional anterior son:

...

donde la elipsis indica un modelo implícito que representa las posibilidades en las que la cláusula si es falsa. Por tanto, consideramos que cualquier situación en la que la cláusula si es verdadera hace que el condicional sea verdadero o falso dependiendo de si la situación encaja o no con el modelo mental explícito. Y consideramos cualquier situación en la que la cláusula si es falsa como irrelevante para la verdad o falsedad del condicional, porque la situación corresponde a un modelo sin contenido explícito (simbolizado por la elipsis).

Consideremos de nuevo la inferencia de Chernobyl del tipo de la que planteé a los estudiantes de ingeniería:

Si en el tablero hay un círculo entonces en el tablero hay un triángulo.

En el tablero no hay un triángulo.

¿Qué se sigue, si se sigue algo?

Si razonamos utilizando modelos mentales del condicional, la premisa categórica de que no hay un triángulo elimina el modelo explícito del condicional y solo deja el modelo sin contenido explícito. Por tanto, parece que no se sigue nada de

las premisas, que es lo que dijeron muchos de los estudiantes. Una forma de alcanzar la conclusión válida de que no hay un círculo es desarrollar los modelos del condicional para incluir la posibilidad:

No hay un círculo y no hay un triángulo.

Otra forma es hacer la suposición de que en el tablero hay un círculo. Se sigue del modelo explícito del condicional que en el tablero hay un triángulo. Esta conclusión contradice la premisa categórica, y por tanto una inferencia posterior basada en modelos da lugar a la negación de la proposición. Una forma más es pensar en el caso que el condicional descarta, es decir, su antimodelo en el que hay un círculo y no un triángulo e inferir del hecho de que no hay un triángulo que no puede haber un círculo. Ninguno de estos modelos es sencillo, y por eso es por lo que la inferencia de Chernobyl es difícil. En un condicional, la cláusula si invoca a menudo el razonamiento partiendo de la suposición de que es cierta y algunos psicólogos han sostenido que a partir de los condicionales razonamos siempre de forma hipotética. Sin embargo, algunas cláusulas si afirman hechos, p. ej., “Si hace este calor ahora, va a hacer aún más calor en julio”.

* * *

Otro misterio sobre los condicionales es raíz de una larga controversia. Se refiere a su interpretación lógica, y el misterio es que una premisa categórica como por ejemplo:

Hay un triángulo.

implica cualquier condicional sea cual sea con tal de que la premisa categórica

sea también la cláusula entonces, p. ej.:

Si hay un círculo entonces hay un triángulo.

La razón es que la interpretación lógica de un condicional puede ser falsa solo cuando es falsa la cláusula entonces, pero la premisa categórica excluye esta posibilidad. Una premisa categórica como:

No hay un círculo.

implica también cualquier condicional sea cual sea, incluido el anterior, con tal de que la premisa categórica niegue su cláusula si. La razón es que la interpretación lógica de un condicional no puede ser falsa cuando la cláusula si del condicional es falsa. Estas inferencias son tan extrañas que algunos autores las han apodado “paradojas”, y rechazan aceptar cualquier teoría de los condicionales que las permita.

Las dos paradojas son similares a inferencias basadas en disyunciones incluyentes:

Hay un triángulo.

Por tanto, no hay un círculo o hay un triángulo.

y:

No hay un círculo.

Por tanto, no hay un círculo o hay un triángulo.

También nosotros rehusamos hacer estas inferencias. Lo que tienen en común con las paradojas condicionales es que malgastan información añadiendo posibilidades alternativas a las que se incluyen en sus premisas categóricas. Ningún teórico, sin embargo, ha rechazado el significado lógico de las disyunciones en base a que dan lugar a conclusiones quasiparadójicas. La teoría de modelos, como verá usted enseguida, predice que las apariciones de algunas paradojas condicionales son más aceptables que las de otras.

* * *

El significado y el conocimiento pueden modular la interpretación de las conectivas proposicionales. Un efecto es bloquear la construcción de modelos. El ejemplo del principio: “Si condujera el coche entonces lo estrellaría”, ilustra este efecto, porque el conocimiento bloquea la construcción de un modelo en el que él no conducía el coche, y aun así lo estrelló. Sabemos que para estrellar un coche una persona tiene que estarlo conduciendo. Incluso un ingeniero que está controlando una prueba de colisión tiene que estar conduciendo el coche para estrellarlo, si bien es cierto que a distancia. Un ejemplo en el que se bloquea un tipo diferente de modelo es este:

Si metes el dedo en la piscina entonces el agua está caliente.

Uno sabe que meter el dedo en la piscina no basta por sí mismo para hacer que el agua esté caliente. Está caliente independientemente de que uno meta el dedo.

Por tanto, la modulación bloquea la posibilidad en la que uno no mete un dedo y el agua no está caliente. El condicional significa, en realidad, el agua está caliente, como descubrirás si metes un dedo en la piscina. Los variados efectos de la modulación se han corroborado en experimentos en los que los participantes enumeraban lo que es posible para condicionales con diferentes contenidos. Las interpretaciones a su vez tienen efectos predecibles sobre el razonamiento.

La ironía puede bloquear dos de los tres modelos de un condicional lógico. Por ejemplo,

Si él está aquí entonces me tiraré al lago

significa que él no está aquí (y, como usted sabe, yo no me tiraré al lago). El condicional:

Si me llamo Phil entonces él está aquí

significa que él está aquí (y, como usted sabe, me llamo Phil).

Un segundo efecto de la modulación es añadir información a los modelos. Puede introducir diversas relaciones entre los sucesos a los que se refiere el condicional. En el condicional “Si puso el libro en la estantería entonces se cayó”, el primer suceso ocurrió antes que el segundo, y el libro acabó en una posición inferior a la que tenía en la estantería. Las relaciones temporales exigen modelos de posibilidades en momentos diferentes. Pueden estar organizados de forma que un modelo aparece después que otro en una secuencia mental de los hechos, o de una forma estática similar a un diagrama o gráfico.

Otra modulación de este tipo añade un mecanismo causal putativo a una interpretación:

Si levantas el interruptor entonces la bombilla se enciende.

Dado que tenemos el conocimiento adecuado, tendemos a inferir que el interruptor forma parte de un circuito eléctrico, que al levantarlo cierra el circuito para hacer que fluya la electricidad a través del circuito y por tanto a través del filamento de la bombilla, lo que a su vez hace que este emita luz. Como muestra el ejemplo, el mecanismo en sí mismo está formado por relaciones causales a menor escala.

Un tercer efecto de la modulación es ayudar a construir modelos completos, es decir, los que enumeran tanto lo que es verdadero como lo que es falso respecto a las posibilidades consistentes con las proposiciones. Por ejemplo, la premisa:

Si Bill está en Río de Janeiro entonces está en Brasil

expresa conocimiento general sobre una relación geográfica. Sabemos que Río está en Brasil. Por tanto, nos damos cuenta de que si Bill no está en Brasil, no puede estar en Río. Un corolario es que deberíamos hacer la inferencia de Chernobyl con facilidad:

Si Bill está en Río de Janeiro entonces está en Brasil.

Bill no está en Brasil.

Por tanto, no está en Río de Janeiro.

Por el contrario, las exclusiones espaciales deberían inhibir este tipo de inferencia, y por tanto dadas las premisas:

Si Bill está en Brasil entonces no está en Río de Janeiro.

Bill está en Río de Janeiro.

deberíamos ser reacios a la conclusión:

Bill no está en Brasil.

Sabemos que si Bill está en Río entonces debe estar en Brasil. Con contenidos familiares de este tipo, los participantes en un experimento extrajeron muchas más conclusiones con las inclusiones espaciales (noventa y dos por ciento) que con las exclusiones espaciales (treinta y cuatro por ciento). Estos efectos se producen con modulaciones que dependen del conocimiento general o del significado, y afectan también a otros tipos de razonamiento condicional.

* * *

La modulación resuelve otros misterios famosos. Ítem: la referencia en los condicionales. La referencia puede ser tan problemática para los análisis lógicos que ciertos ejemplos se conocen en algunos ámbitos como oraciones del “burrito” [“donkey” sentences] a partir del ejemplo original del problema. Una versión simplificada es esta:

Si Joe tiene un burro entonces le pega.

¿A qué se refiere aquí “le”? Si Joe tiene un burro, el pronombre se refiere al burro, pero si no tiene un burro, el pronombre no se refiere a nada. El condicional solo es compatible con dos posibilidades:

Joe tiene un burro y le pega.

Joe no tiene un burro.

Ítem: de acuerdo con el significado lógico de los condicionales es válida este tipo de inferencia, conocida como “reforzamiento del antecedente”:

Si uno frota adecuadamente una cerilla entonces se enciende.

Por tanto, si una cerilla está empapada y se la frota adecuadamente entonces se enciende.

Acertijos análogos se producen con las proposiciones disyuntivas, p. ej.:

Se pone azúcar en el café o no sabe dulce.

Por tanto, se pone azúcar o gasóil en el café, o no sabe dulce.

Se puede pensar que el problema surge porque las premisas de estos ejemplos no son ciertas, y que la solución es reformularlas, esmerándose en establecer todas las condiciones relevantes, p. ej.:

Si una cerilla no se ha mojado con agua, está químicamente intacta, está en una atmósfera adecuada que contiene oxígeno... y se frota adecuadamente entonces se enciende.

¿Se puede garantizar que se establecen todas las condiciones relevantes? Probablemente, no. Por tanto, esta solución puede degenerar en un círculo vicioso:

Si una cerilla está en un estado en que se enciende cuando se la frota adecuadamente y se la frota adecuadamente entonces se enciende.

Si a uno le gusta este tipo de solución, entonces esta solución es del tipo que a uno le gusta. Cuando razonamos en la vida real, casi siempre razonamos en ausencia de información completa. Llegamos a una conclusión precipitada que la realidad puede invalidar. Sin embargo, nuestras conclusiones a menudo son ciertas. Y por eso nuestra creencia, pongamos por caso, de que cuando las cerillas se frotan adecuadamente se encienden es una idealización que es un trampolín útil para hacer inferencias.

Cuando este misterio atrajo por primera vez la atención de los profesionales de la inteligencia artificial –los científicos que tratan de escribir programas de ordenador para hacer cosas inteligentes– una de las figuras en ese campo, Marvin Minsky, introdujo la idea de una asunción por defecto. Es decir, una asunción que puede hacerse dado que no existen pruebas para rechazarla. Si una cerilla se frota adecuadamente y no hay información en contra, entonces, por defecto, se enciende. Mojar una cerilla es una forma de invalidar esta asunción

por defecto. No obstante, un enfoque más psicológico del problema es admitir que el conocimiento puede modular la interpretación de las conectivas proposicionales. Considere de nuevo la inferencia:

Si uno frota adecuadamente una cerilla entonces se enciende.

Por tanto, si una cerilla está empapada y uno la frota adecuadamente entonces se enciende.

La interpretación lógica de la conclusión da lugar a un modelo de una posibilidad en la que la cerilla está empapada, se frota adecuadamente y se enciende. Sin embargo, la proposición de que la cerilla está empapada desencadena nuestro conocimiento de que las cerillas empapadas no se encienden. El conocimiento se impone al significado de las oraciones, y por tanto este modelo excluye la cláusula “entonces” para dar lugar a una posibilidad alternativa: la cerilla está empapada, se frota adecuadamente y no se enciende. Esta posibilidad es un contraejemplo para la premisa condicional, y por tanto la conclusión no se sigue de la premisa.

Escribí un programa de ordenador que utiliza una “base de datos” de conocimiento de esa misma forma. Si al programa se le dan las premisas:

Si se frota la cerilla entonces la cerilla se enciende.

La cerilla está mojada y se frota.

entonces la cláusula que dice que la cerilla está mojada desencadena en la base de datos una posibilidad que se ajusta a ella según la cual la cerilla no se

enciende. Este conocimiento tiene preferencia, y bloquea la conclusión de que la cerilla se enciende.

Ítem: las “paradojas” de la interpretación lógica de los condicionales. Algunos autores, como he mencionado, rechazan las paradojas junto con cualquier teoría que las justifique. Por el contrario, la teoría de modelos predice que deberían ser aceptables en algunos casos. Un caso es aquel en que un condicional se modula de forma que no hay duda sobre que la cláusula si es falsa, p. ej.:

El experimento no funcionará.

Por tanto, si el experimento funciona entonces me multiplicaré por cero[1].

El otro tipo de paradojas debería ser aceptable cuando el condicional se modula de forma que no hay duda sobre la veracidad de la cláusula entonces, p. ej.:

No jugaron al fútbol.

Por tanto, si jugaron a algo entonces no jugaron al fútbol.

Estas inferencias son verosímiles, y por tanto las teorías que rechazan todos los casos de paradojas pueden estar equivocadas.

Ítem: conjunciones y disyunciones de condicionales. Consideremos esta premisa:

Si hay un círculo entonces hay un triángulo y si hay un rombo entonces hay una

estrella.

¿Cuáles son las posibilidades compatibles con ella? Tendemos a enumerar tres posibilidades: una en la que hay un círculo y un triángulo, otra en la que hay un rombo y una estrella y una en la que están presentes las cuatro formas. Ahora, consideremos esta premisa alternativa:

Si hay un círculo entonces hay un triángulo o si hay un rombo entonces hay una estrella.

¿Cuáles son las posibilidades compatibles con ella? Tendemos a enumerar las mismas tres posibilidades. Existe una diferencia de significado entre y y o, por tanto, ¿por qué enumeramos las mismas posibilidades? La respuesta es que las dos oraciones tienen los mismos modelos mentales, porque los modelos mentales no representan proposiciones que son falsas en las posibilidades. Sin embargo, los modelos completos de los dos asertos, que tienen en cuenta la falsedad de las oraciones, dan lugar a nueve posibilidades para la conjunción y quince posibilidades para la disyunción. Nadie en ningún experimento ha llegado nunca a acercarse ni de lejos a la enumeración de estas posibilidades.

* * *

¿Es verdadero o falso este condicional?

Si la constitución de los Estados Unidos no se enmienda para permitir que las personas nacidas fuera del país puedan ser elegidas presidente, entonces Arnold Schwarzenegger será un día presidente de Austria.

No podemos decirlo de momento. No obstante, podemos creer en mayor o menor grado en el condicional, y podemos aportar pruebas para apoyar nuestra posición. Podemos pensar internamente siguiendo esta línea:

Bien, es posible. Mr Schwarzenegger es un político bastante hábil; no es un secreto que le gustaría presentarse a la Presidencia de Estados Unidos, pero eso exigiría una enmienda constitucional. Por tanto, si no se aprobara una enmienda, y dudo bastante de que se hiciera, entonces él podría volver a su país de origen, y ser allí un político de éxito. Por otra parte, hay muchos imponderables y por tanto no es ni mucho menos seguro.

No podemos juzgar con seguridad si el condicional es verdadero o falso. Pero lo que desconcierta a los psicólogos son los mecanismos que establecen la fuerza con que creemos en él.

Una opinión con gran influencia se remonta al lógico inglés Frank Ramsey. Planteó esta cuestión en una nota a pie de página en un artículo, una de las notas a pie de página más citadas de la filosofía:

Si dos personas están discutiendo "¿Si p entonces q?" y ambos dudan de p, están añadiendo hipotéticamente p a su bagaje de conocimiento y discutiendo sobre q a partir de esa base; de modo que en un sentido 'Si p, q' y 'Si p, no q' son contradictorios. Podemos decir que están estableciendo su grado de creencia en q dado p. Si p resulta ser falso, estos grados de creencia se vuelven nulos.

En la versión de Ramsey, añadimos la cláusula si a nuestras creencias y luego determinamos si creemos en la cláusula entonces. Y ¿cómo hacemos eso? No es algo obvio. Podríamos utilizar nuestro conocimiento para construir una cadena de implicaciones, quizá de naturaleza causal, como se ilustraba antes en el ejemplo sobre Mr. Schwarzenegger. Varios filósofos modernos, sin embargo, han

transformado el test de Ramsey en una teoría, no respecto a las creencias, sino respecto al significado de los condicionales. Consideran que un condicional es verdadero si, en cualquier "mundo posible" en el que la cláusula si es verdadera, y que está tan cerca como sea posible del mundo real, la cláusula entonces también es verdadera. Esta teoría permite que la cláusula si sea falsa, y por tanto un mundo posible en el que se cumple la cláusula si y se parece al mundo real es una idealización filosófica. La teoría no se puede poner en práctica para determinar si un condicional real es verdadero o falso. Lo que es más preocupante es que cuando sabemos que un condicional es verdadero, p. ej., Si mi próximo artículo es aceptado para su publicación estaré satisfecho, no podemos establecer que sea cierto según los principios de esta explicación.

Para volver a los grados de creencia, la nota a pie de página de Ramsey es verosímil, aunque los psicólogos no tengan un modelo de trabajo para ella. Pero parece que no es la historia completa. Consideremos esta afirmación:

Si Bacon escribió Hamlet, entonces no lo hizo Shakespeare.

Lo creo, incluso aunque piense que la cláusula si –que Bacon escribió Hamlet– es falsa. En este caso no se aplica el test de Ramsey, pero siguiendo el ejemplo de sus idealizaciones filosóficas, podríamos decir que añado la cláusula si a mis creencias, las ajusto para acomodar a este intruso y entonces calculo si creo en la cláusula entonces. En realidad, para llegar a creer que el condicional es verdadero, no creo ni por un momento que Bacon escribiera Hamlet. Simplemente pienso:

Bacon y Shakespeare fueron dos individuos diferentes.

Hamlet fue la obra de un individuo.

Ergo, si Bacon escribió Hamlet, Shakespeare no lo hizo.

* * *

Con frecuencia razonamos sobre lo que no sucedió. Entramos en un mundo imaginario en el que un suceso que una vez fue posible no llegó a producirse. Algunas veces nuestros pensamientos se refieren a sucesos triviales –si los vieneses hubieran tenido tres piernas entonces habrían andado a ritmo de vals–. Otras veces se refieren a asuntos de vida o muerte. El historiador militar Max Hastings, por ejemplo, escribió sobre la debacle de Arnhem en la Segunda Guerra Mundial:

Muchas descripciones... se han concentrado en los "podría haber sido" del fallo británico en Arnhem. Pero parece al menos igual de relevante examinar los de Nijmegen. Si elementos del 82 aerotransportado hubieran aterrizado más cerca del puente, y si la amplia fuerza aliada de cazabombarderos se hubiera utilizado para bloquear a los vehículos acorazados alemanes que irrumpieron en la batalla a lo largo de las rutas abiertas holandesas, el cruce se podría haber tomado en el primer día.

El extraordinario libro de Ruth Byrne The Rational Imagination informa de la importante investigación realizada por ella y sus colegas sobre este tema de los condicionales "contrafactuales". Su punto de partida es que al imaginar alternativas a la realidad, los sujetos tienden a alterar los mismos tipos de cosas. Sostiene que nuestro pensamiento sobre las posibilidades contrafactuales está dominado por un pequeño conjunto de principios tácitos. Tendemos a deshacer las acciones voluntarias, y aquellas que son inaceptables pero que tuvieron una influencia causal en el resultado. A corto plazo, tendemos a lamentar lo que hicimos más que lo que no hicimos, pero a largo plazo –pongamos por caso, cuando pensamos en nuestra vida como un todo–, lamentamos más las acciones que no realizamos.

Un condicional, como por ejemplo:

Si él hubiera salido de casa antes de las 6 de la mañana entonces no se habría encontrado un atasco de tráfico

se puede utilizar para hacer una afirmación contrafactual. Y cuando lo entendemos de esta forma, consideramos dos posibilidades. Una posibilidad es lo que presupone que ha sucedido la persona que habla, es decir, la realidad:

Él no salió de casa antes de las 6 de la mañana. Se encontró un atasco de tráfico.

La otra es una posibilidad contrafactual, es decir, es algo que en un determinado momento fue una posibilidad real, pero que en realidad no sucedió:

Él salió de casa antes de las 6 de la mañana. No se encontró un atasco de tráfico.

Como Byrne y sus colegas han mostrado, una consecuencia de este tipo de representación es que nos hace más fácil extraer la inferencia tipo Chernobyl. Es decir, a partir de la última premisa:

En realidad, se encontró un atasco de tráfico

tendemos a sacar la conclusión:

Él no salió de casa antes de las 6 de la mañana.

Estas inferencias son más fáciles que las correspondientes basadas en condicionales ordinarios, porque hemos construido un modelo explícito que corresponde a la falsedad de la cláusula si.

* * *

Las diferencias con otras versiones de los condicionales ponen más de relieve la teoría de modelos. Una de las versiones se ajusta a los juicios ingenuos sobre la verdad y falsedad que describí anteriormente. Postula que un condicional, como por ejemplo:

Si en el tablero hay un círculo entonces en el tablero hay un triángulo

es verdadero o falso solo cuando la cláusula si es verdadera. Si la cláusula si es falsa, el condicional no es ni verdadero ni falso. Como dicen los filósofos, no tiene valor de verdad. Una razón para dudar de la teoría surge de un estudio que llevaron a cabo Tom Ormerod y sus colegas. Mostraron que inferimos disyunciones a partir de condicionales y condicionales a partir de disyunciones. Sin embargo, nadie supone que una disyunción no tiene valor de verdad porque una de sus cláusulas es falsa. Otro golpe a esta teoría viene de los condicionales que sabemos que son verdaderos. Consideremos el condicional:

Si se van a destinar fondos para la próxima petición de subvención de Phil entonces él estará encantado.

Es cierto (créame). Pero si los condicionales no tienen valor de verdad cuando sus cláusulas si son falsas, se deduce que la cláusula si de este condicional debe ser verdadera:

Se van a destinar fondos para la próxima petición de subvención de Phil.

Esta milagrosa inferencia se sigue de la teoría debido a que la única forma de que un condicional sea verdadero es que la cláusula si, contenida en él, sea verdadera. ¡Qué maravillosa receta para el pensamiento ilusorio! Es demasiado bueno para ser cierto, y demasiado bueno para la teoría respecto a la desaparición del valor de verdad no tener valor de verdad. Es falsa.

Otra versión de los condicionales tiene que ver con cuándo está justificado darles por válidos. Según esta explicación, está justificado dar por válido el condicional:

Si hay un círculo entonces hay un triángulo.

cuando la probabilidad de:

Hay un triángulo dado que hay un círculo

es mayor que la probabilidad de:

No hay un triángulo dado que hay un círculo.

El atractivo de esta versión reside, sospecho, en el efecto migración que describí anteriormente. Alguien nos pregunta:

¿Cuál es la probabilidad de que si hay un círculo haya un triángulo?

La migración nos conduce a interpretar la pregunta como equivalente a:

Si hay un círculo, entonces, ¿cuál es la probabilidad de que haya un triángulo?

Este constructo convierte la pregunta original en una petición directa de una probabilidad condicional. Es difícil resistir a este constructo debido a que está justificado por la gramática de las cláusulas subordinadas. Sin embargo, induce a error, debido a la equivalencia entre este tipo de condicionales y las disyunciones y a que la probabilidad de las disyunciones depende, no de una probabilidad condicional, sino de en qué posibilidades se cumple.

* * *

¿Cuáles son las soluciones al misterio de los condicionales? Dependen de tres puntos esenciales:

En primer lugar, entendemos que un condicional se refiere a un conjunto de posibilidades, pero rara vez pensamos en todas ellas. La inferencia tipo Chernobyl es difícil porque tendemos a no representar para el condicional las posibilidades en las que la cláusula si es falsa.

En segundo lugar, un condicional con frecuencia suscita, entre la situación descrita en la cláusula si y la situación descrita en la cláusula entonces,

relaciones espaciales, temporales o de algún otro tipo. Lo que captamos de su contenido, nuestro conocimiento, e incluso nuestra imaginación, contribuyen conjuntamente a interpretaciones que dan lugar a esas relaciones. Por tanto, los condicionales tienen un número indeterminado de tipos de interpretación diferentes.

En tercer lugar, puede producirse la interpretación lógica de un condicional, y la modulación puede incluso hacer más verosímiles algunas de sus consecuencias paradójicas. Pero el proceso de interpretación trasciende la lógica del cálculo proposicional, porque siempre debemos estar a la caza de los efectos de la modulación.

La complejidad de los condicionales tiene causas sencillas. Sus características camaleónicas surgen de las interacciones entre un conjunto de componentes elementales: su gramática, su significado lógico central, su representación en modelos y su modulación debida al contenido y el conocimiento. Lo que sigue siendo un misterio es cómo determinamos nuestra creencia en un condicional. Pregúntese a sí mismo una última vez si cree en la proposición:

Si la constitución de los Estados Unidos no se enmienda para permitir que las personas nacidas fuera del país puedan ser elegidas presidente, entonces Arnold Schwarzenegger será un día presidente de Austria.

Puede que haya pensado en lo que sucedería si la constitución no fuera enmendada, dadas las ambiciones de Mr Schwarzenegger, y luego haya concedido que la cláusula entonces es posible. Ramsey tenía razón. Pero, ¿no pensó usted también –solo por un momento– sobre lo que podría suceder si la constitución de los Estados Unidos fuera enmendada? ¿Arnold como presidente de los Estados Unidos? ¿O se presenta, pierde y regresa a su país de origen? Si es así, Ramsey nos contó solo parte de la historia. Ninguna teoría psicológica existente describe los mecanismos que nos capacitan para asignar un número a nuestro grado de creencia en un condicional. Esa peliaguda palabrita, “si”, va a seguir molestándonos una temporada.

Notas al pie

[1] N. de la tr.: La expresión inglesa “eat my pants” es una frase célebre del personaje televisivo Bart Simpson que ha sido traducida en España como “Multiplícate por cero” y en Latinoamérica como “Cómete mis calzoncillos”.

CAPÍTULO 22

Causas y obligaciones

En un capítulo anterior describí cómo funciona una máquina de coser. La aguja hace un agujero en la prenda, y empuja a través de él un lazo de hilo. La rotación de la bobina permite que la muesca que hay en ella enganche el lazo, y causa que al tirar de él este rodee la bobina, que contiene un hilo inferior independiente. Un extremo de este hilo sale de las prendas –ya se ha cosido a ellas, pero el otro extremo del hilo está libre– está enrollado en la bobina. Cuando el lazo sale por el otro lado de la bobina, la aguja se está moviendo hacia arriba y por tanto arrastra al lazo que contiene el hilo inferior desde la bobina hacia arriba, de manera que tira de él contra la parte inferior de las prendas. Y ya puede repetirse el ciclo completo.

Esta explicación describe los mecanismos causales de una máquina de coser moderna – descendiendo hasta el nivel de simples acciones mecánicas–. Si vuelve a fijarse en la descripción, verá que contiene una serie de verbos que expresan causación, sea de manera explícita, como en el caso de hacer y causar, o implícita, como en el caso de empujar y tirar de. Hume escribió: “Todos los razonamientos que se refieren a hechos parecen basarse en la relación de causa y efecto.” Esta afirmación parece verosímil, pero no todo el mundo está de acuerdo, y algunos autores han sostenido incluso que la causación no existe. El concepto, dicen, es incoherente. Sin embargo, forma parte de tantos verbos, como por ejemplo tirar de y empujar, que su incoherencia parece improbable.

Existe un paralelismo entre las relaciones causales y las relaciones deónticas, es decir, las que conciernen a acciones obligatorias y lícitas. Si un suceso causa otro, entonces el que se produzca el primer suceso hace que el segundo sea una necesidad física. Si una acción obliga a otra, entonces el que se produzca la primera acción hace de la segunda acción una necesidad moral. La necesidad física y la moral se diferencian en el significado. Una necesidad física está obligada a producirse, pero una necesidad moral no está obligada a producirse –de vez en cuando hay individuos que no cumplen con sus obligaciones–. Una

oración como “Debe suceder” tiene un significado ambiguo entre ambos. El lenguaje de la causación es cercano al lenguaje de la obligación, aunque no idéntico a él.

¿Qué significan las proposiciones causales y deónticas? Estos son los misterios que espero aclarar con el fin de explicar nuestro razonamiento respecto a ellas. Las causas, que son el tema de la primera mitad del capítulo, pueden ser acciones, sucesos o meras situaciones – como escribió Benjamin Franklin, “a falta de un clavo se perdió el zapato”–. Las obligaciones, que son el tema de la segunda mitad del capítulo, se refieren a situaciones posibles y lícitas.

* * *

Nuestras creencias cotidianas sobre la causación algunas veces son incoherentes. Asumimos que iniciamos una cadena causal. Servimos en tenis; nuestro oponente devuelve la bola; y jugamos de aca para allá hasta que acaba el peloteo. Volvemos a servir, etc. Cada servicio es el inicio de una cadena causal, que finaliza cuando se decide el punto. De manera similar, un experimento psicológico típico tiene una serie de “pruebas”. En algunas pruebas, la experimentadora nos presenta, digamos, una taza de té en la que echó la leche en primer lugar, y en otras pruebas, nos presenta una taza de té en la que ha echado la leche en segundo lugar, probamos el té y tratamos de identificar si la leche se puso en primer o en segundo lugar. Cada prueba es una cadena causal independiente, y la experimentadora está interesada en averiguar si podemos decir la diferencia entre los dos tipos de té. También la naturaleza puede intervenir para iniciar y terminar una cadena causal. Un niño está volando una cometa, cuando una ráfaga de viento golpea la cuerda y la cometa cae al suelo. La noción de que una intervención inicia una cadena causal parece sensata. Por el contrario, sin embargo, con frecuencia asumimos que cada suceso tiene una causa. La pantalla de nuestro ordenador parpadea y se apaga. Inferimos que algo ha ido mal: quizá, la batería se ha acabado y necesita ser recargada. Pero, si cada suceso tiene una causa, entonces la propia acción que parece empezar una cadena causal tiene una causa, y por tanto no es el inicio de la cadena. Esta opinión nos retrotrae a los orígenes del universo, que es la causa sin causa que inició todas las cadenas causales. Sin embargo, si un jugador de tenis, un

psicólogo y una ráfaga de viento pueden todos ellos iniciar cadenas causales, entonces estamos considerando cada una de las acciones iniciadoras como sin causa. Tenemos que elegir entre una cosa y otra: o bien podemos iniciar una cadena causal o bien cada suceso tiene una causa. No obstante, estas son asunciones tácitas sobre la causación, y no deberían confundirse con el significado del término, que analizaré dentro de un momento. Cuando pensamos sobre el comienzo de una cadena causal, prescindimos por simplicidad de sus causas previas.

Algunas veces se dice que la iniciación de una cadena causal, o abreviando una intervención, tiene su propia lógica especial. Pero, como ejemplo, consideremos la afirmación causal de que comer demasiado causa indigestión. Contando con que es cierta, si observáramos que Phil no tenía indigestión, entonces inferiríamos que no había comido demasiado. Pero, supongamos que Phil hubiera hecho una intervención: tomó una pastilla que prevenía la indigestión. Ahora, no seguiríamos infiriendo de su falta de indigestión que no había comido demasiado. Su intervención inició una nueva cadena causal que anula los efectos de comer demasiado. No se requiere una lógica especial, sino solo nuestra capacidad para entender las premisas:

Comer demasiado causa indigestión.

Tomar una pastilla contra la indigestión impide la indigestión.

y darnos cuenta –como en el caso de la modulación– de que la segunda premisa tiene prioridad sobre la primera.

* * *

Una proposición como por ejemplo:

El uso de fertilizante hará que las plantas crezcan

hace una afirmación sobre una relación causal determinada. Es equivalente a la proposición condicional:

Si se usa el fertilizante entonces las plantas crecerán.

En otras palabras, el significado de ambas oraciones es consistente con las mismas tres posibilidades: se usa el fertilizante y las plantas crecen, no se usa el fertilizante y las plantas no crecen y no se usa el fertilizante pero las plantas crecen de cualquier modo. Lo que es imposible de acuerdo con ambas proposiciones es que se use el fertilizante y las plantas no crezcan. Por tanto, use el fertilizante, y solo hay una posibilidad: las plantas deben crecer.

Cuando se conoce el resultado, la afirmación es directa:

El uso del fertilizante hizo crecer las plantas.

El correspondiente condicional es “contrafactual” en el sentido que discutí en el capítulo anterior:

Si no se hubiera utilizado el fertilizante entonces las plantas no habrían crecido o podrían no haber crecido.

Los hechos son que se usó el fertilizante y las plantas crecieron, mientras que las

posibilidades contrafactuales son que no se usó el fertilizante y las plantas no crecieron o crecieron por alguna otra razón. Lo que no fue posible es que se utilizara el fertilizante y las plantas no crecieran.

Este análisis no es original. Hume hizo más o menos la misma afirmación cuando escribió: “Podemos definir una causa como un objeto al que sigue otro, siendo así que todos los objetos similares al primero son seguidos por objetos similares al segundo. O en otras palabras, si el primer objeto no hubiera estado, el segundo nunca habría existido. ” ¿Por qué no acaba aquí entonces la presente descripción de la causación? La respuesta en una palabra es: complicaciones.

* * *

El hecho de que las afirmaciones causales se puedan parafrasear en forma de condicionales no implica que todos los condicionales expresen afirmaciones causales. La razón es que la modulación permite a los condicionales expresar una diversidad de tipos diferentes de proposición. Podría usted suponer que el análisis anterior utiliza los condicionales en su significado lógico según el cual son compatibles con tres posibilidades. Una restricción normal, sin embargo, es que un efecto no precede a su causa. Esta restricción temporal permite a una causa ser simultánea con su efecto –el apretar el tubo de dentífrico se produce al mismo tiempo que la salida de la pasta del tubo–. Hume en su Treatise on Human Nature sostiene que la causa precede al efecto y rechaza la simultaneidad. No obstante, declaró que el asunto no tenía gran importancia. Sin embargo, una interpretación causal de una oración como por ejemplo la Luna causa las mareas, desde el punto de vista de la teoría de la gravitación de Newton, exige una acción instantánea a distancia. La mecánica cuántica moderna también exige la acción a distancia. Y los físicos y los filósofos van un paso más allá cuando discuten el viaje en el tiempo y afirman que un suceso actual causó un suceso pasado. Sin embargo, la restricción temporal parece ser cierta en las proposiciones factuales.

Cuando vemos una bola de billar que golpea a otra y la desplaza, la correcta sincronización da lugar a una impresión irresistible según la cual la primera bola causó el movimiento de la segunda. Pero nuestro concepto de causación es

mucho más amplio y más abstracto que las observaciones sobre bolas de billar que golpean a otras. Una causa no necesita chocar con su efecto. Podemos afirmar, por ejemplo, que la mala educación de Pat causó que Viv perdiera los nervios, sin tener que imaginar una cadena de sucesos contiguos en el espacio y el tiempo que lleven de la causa al efecto. Puede existir esa cadena, es decir, un mecanismo causal que los enlace, pero, como veremos, su existencia no es parte del significado cotidiano de las proposiciones causales.

* * *

La afirmación de que una causa causa un efecto es compatible con tres posibilidades restringidas temporalmente:

causa efecto

no causa efecto

no causa no efecto

Dada la causa, como sostuvo Kant, el efecto es por tanto necesario. Pero existe otro tipo de relación causal en la vida real. La televisión pública en Estados Unidos trasmite con frecuencia el mensaje, “Este programa ha sido posible gracias a espectadores como usted”. Significa, no que los espectadores hicieran el programa –es decir, causaran el que existiera– , sino que hicieron posible que se hiciera. Es una relación por tanto en la que un estado hace posible que se produzca otro. Por consiguiente, la proposición:

Mantener esta dieta le hará posible (o permitirá) perder peso.

significa, no que la dieta causará que se pierda peso, sino que hará posible que se pierda peso. También se podría perder peso aunque no se siguiera la dieta. En otras palabras, todos los casos son posibles: se podría perder peso o no, independientemente de que se guardara la dieta o no. Sin embargo, a menudo, queremos decir que no se perderá peso a menos que se siga la dieta. En este sentido más estricto, la proposición ofrece más información. Solo es compatible con tres posibilidades, pero son diferentes de las de una afirmación causal:

posibilitador	efecto
posibilitador	no efecto
no posibilitador	no efecto

Con una causa el efecto es necesario, con un posibilitador es posible.

Los dos conceptos pueden definirse en términos recíprocos. Si la luz causa que suene la alarma entonces la luz no permite que la alarma no suene. De manera similar, si la luz permite que la alarma suene entonces la luz no causa que la alarma no suene, es decir, no le impide sonar. Estas definiciones funcionan porque escondidos en necesidad y posibilidad se apela a “todos” y “algunos”: un suceso es necesario si se cumple en todos los estados posibles, y es posible si se cumple en al menos uno de esos estados. Y se puede definir “todos” y “algunos” de manera recíproca: todos los A son B significa que no se da el caso de que algunos A no sean B, y algunos A son B significa que no se da el caso de que todos los A no sean B.

La proposición según la cual una dieta causa el que se pierda peso es una

relación causal débil en el sentido de que hay otras formas de poder perder peso. Se puede hacer ejercicio de manera regular, por ejemplo. Algunas veces, sin embargo, una proposición causal es más sólida en el sentido de que reconoce una causa única, como por ejemplo en: beber demasiado alcohol causa que uno se emborrache. El alcohol es suficiente para emborracharse, pero también es necesario: no hay otra forma de emborracharse hasta donde yo sé. Entendemos las proposiciones sobre causas y condiciones posibilitadoras, pero no podemos tener presentes todas las posibilidades consistentes con ellas. La teoría de modelos postula que los modelos mentales más destacados son los mismos para ambos tipos de proposición –aquellos en los que aparecen la causa o la condición posibilitadora y el efecto–. Una consecuencia es que tenemos dificultades para distinguir entre las dos. De manera similar, muchos teóricos empezando por el filósofo del siglo diecinueve John Stuart Mill, han negado que exista ninguna diferencia de significado entre causas y condiciones posibilitadoras. Las versiones modernas ofrecen todo tipo de distinciones entre ellas: la causa es un estado no habitual y la condición posibilitadora es el estado habitual, la causa es inconstante, mientras que la condición posibilitadora es constante, la causa viola una norma, mientras que la condición posibilitadora no lo hace, etc. ¿Por qué han seguido los teóricos a Mill en este tema? La respuesta, sospecho, es que es difícil mantener presentes las cinco posibilidades diferentes a las que se refieren las descripciones de cómo una causa y un posibilitador trabajan a la vez para dar lugar a un efecto (véase más adelante la Tabla 22.1), y también es demasiado fácil pensar en casos con dos causas juntas.

A pesar de todo, los experimentos han mostrado que somos sensibles en alguna medida a la diferencia entre causas y condiciones posibilitadoras. Tendemos a enumerar posibilidades diferentes para ambas, en línea con la teoría. Consideramos que una expresión con la forma A causará B excluye como imposible un caso en el que se produzca A pero no se produzca B, mientras que consideramos que una expresión con la forma A permitirá B es compatible con esta posibilidad –si excluye algo, es la posibilidad de que se produzca B en ausencia de A–. Por tanto, a partir de los dos tipos de afirmación deberíamos extraer conclusiones deductivas diferentes. Y un experimento corroboró también esta predicción. Dadas las premisas:

Comer proteínas causará que ella gane peso.

Comerá proteínas.

los participantes tendían a extraer la conclusión: Ella ganará peso. Pero, como predice la teoría, tendían a no sacar esa conclusión a partir de las premisas:

Comer proteínas le permitirá ganar peso.

Comerá proteínas.

En cambio, dadas las premisas:

Comer proteínas le permitirá ganar peso.

No comerá proteínas.

los participantes tendían a extraer la conclusión: Ella no ganará peso. Pero tendían a no sacar esa conclusión a partir de las premisas:

Comer proteínas causará que ella gane peso.

No comerá proteínas.

Nos apoyamos en modelos mentales de las relaciones causales, que representan las posibilidades más destacadas, lo mismo que en el caso de los condicionales. Una consecuencia es que deberíamos sucumbir a inferencias ilusorias para ciertas premisas causales. Una oración con la forma A causa B tiene un modelo mental explícito que representa el caso en el que se producen A y B, y una oración con la forma A evita B tiene un modelo mental explícito que representa el caso en el que se produce A y no se produce B. Por tanto, premisas de la forma:

A evita B

B causa C

dan lugar a modelos mentales explícitos de dos posibilidades. En una posibilidad, se produce A, y no B y C . En la otra posibilidad, no se produce A, pero sí B y C. Por tanto, deberíamos inferir que A evita C. En un experimento que utilizaba contenidos cotidianos adecuados, los participantes realmente hicieron esta inferencia. Pero es una ilusión: C podría tener una causa distinta a B, y por tanto cuando se produce A esta otra causa podría seguir provocando C. El problema de control para esta ilusión tenía la forma:

A causa B.

B previene C.

En este caso se sigue que A evita C, y los participantes también extrajeron esa

conclusión. El experimento examinaba una serie de parejas de problemas de este tipo, y mostró que los participantes sucumbían a las ilusiones, pero eran exactos al razonar sobre sus correspondientes problemas de control.

* * *

En el siglo veinte sucedió algo extraño. Los teóricos propusieron que la causación es una noción probabilística. Algunas afirmaciones generales, como: el fumar causa humo, son verdaderas, pero otras afirmaciones generales, como: el fumar causa cáncer, no son estrictamente verdaderas. No todo el mundo que fuma contrae cáncer. He comprobado que los fumadores están preparados para citar excepciones. El filósofo Hans Reichenbach no rechazaba estas generalizaciones como falsas, sino que postulaba que la causación es probabilística. La proposición de que el fumar causa cáncer, habría dicho, significa que la probabilidad de cáncer en el caso de que una persona fume es mayor que la probabilidad de cáncer en el caso de que una persona no fume. Pero el significado de la causación incluye algo más que esta diferencia entre dos probabilidades condicionales. Consideremos la idea absurda de que el cáncer de pulmón cause que fumemos. La probabilidad de fumar en el caso de una persona que tenga cáncer de pulmón es mayor que la probabilidad de fumar en el caso de una persona que no tenga cáncer de pulmón. Sin embargo, no inferimos que el cáncer de pulmón cause el fumar. A pesar de todo, muchos filósofos han seguido la senda de Reichenbach. Pero esta opinión habría asombrado a Hume, quien consideró que la causación significa una conjunción constante de causa y efecto, y habría asombrado a Kant y a Mill, quienes consideraron que la causación implica una conexión necesaria. Podemos parafrasear una afirmación causal de esta forma: Si una persona fuma entonces debe haber humo. ¿Vamos a asumir que “debe” tiene una interpretación probabilística? La asunción infectaría nuestras obligaciones morales con una peligrosa calificación probabilística. Si usted me presta dinero entonces yo debo devolvérselo. Qué bien le vendría a una persona inmoral que recibe un préstamo que su obligación fuera solo probable.

El método probabilístico podría ser justificable para concepciones científicas de la causación en la mecánica cuántica. Pero es inverosímil como versión de la causación cotidiana. Hay tres líneas de pruebas que están en su contra. Una línea

es que interpretamos que las proposiciones causales significan que ciertos estados son imposibles. Por ejemplo, los experimentos han mostrado que dada la oración: Hacer correr la nueva aplicación causará que el ordenador falle, los participantes tienden a enumerar como imposible el caso en el que se hace correr la nueva aplicación y el ordenador no falla. Este resultado es contrario a la teoría probabilística, que permite que sean posibles todos los casos: lo que importan son sus probabilidades. La utilización de afirmaciones sobre sucesos concretos podría haber disuadido a los sujetos de las interpretaciones probabilísticas, pero una réplica del experimento mostró el mismo efecto para afirmaciones causales generales.

Una segunda línea de pruebas surge de las inferencias que describí con anterioridad. Dadas premisas como: Comer proteínas causará que ella gane peso, y comerá proteínas, tendemos a inferir que ella ganará peso. La inferencia no está garantizada si la primera premisa establece una probabilidad –a menos que la probabilidad sea del 100 por cien–. Pero si asumimos esta probabilidad en el caso de una relación causal que no conocemos, entonces el análisis probabilístico es superfluo.

En la Tabla 22.1 se resume una tercera línea de pruebas. Muestra que para las explicaciones probabilísticas no hay una forma clara de establecer una distinción entre causas y condiciones posibilitadoras. Pero en los significados de las dos relaciones hacemos esta distinción, y la diferencia afecta a su vez a qué conclusiones inferimos. ¿Por qué aceptamos tan a menudo generalizaciones causales para lo que sabemos que son excepciones? Una razón puede ser que las consideramos como idealizaciones útiles, como el anterior ejemplo: frotar adecuadamente una cerilla causa que se encienda. Somos conscientes de que muchas causas de la vida cotidiana solo producen sus efectos si se dan condiciones posibilitadoras y no se dan condiciones inhibidoras desconocidas. Por tanto, cuando afirmamos generalizaciones, como por ejemplo, el fumar causa cáncer, estamos afirmando que el fumar causa cáncer en algunas personas. Es una forma de hablar similar a otros tipos de generalizaciones, como por ejemplo: toda la comida en Italia es buena, o todo el mundo va de rosa esta temporada. Otra razón para nuestras generalizaciones causales es que a menudo las inducimos a partir de datos probabilísticos. Los científicos registran apariciones simultáneas de sucesos, a veces para comprobar conjeturas, a veces para inducir leyes causales. Las probabilidades son útiles para capturar la relación entre sucesos que se producen simultáneamente, de forma que podemos inducir causas a partir de los datos.

Tabla 22.1. Las probabilidades no pueden distinguir entre causas y condiciones posibilitadoras

1. Supongamos que se observan las frecuencias de varios sucesos dependientes de la existencia de luz solar, el uso de fertilizante y el crecimiento de las plantas. Y supongamos también que cada una de las cinco posibilidades se producen con la misma frecuencia, pongamos por caso, veinte veces en cada 100 observaciones:

luz solar	fertilizante	crecimiento	20
luz solar	no fertilizante	crecimiento	20
luz solar	no fertilizante	no crecimiento	20
no luz solar	fertilizante	no crecimiento	20
no luz solar	no fertilizante	no crecimiento	20

2. La probabilidad de crecimiento dada la luz solar (40/60 = 2/3) es mayor que la probabilidad de crecimiento sin luz solar (0/40 = 0).

3. La probabilidad de crecimiento dado el fertilizante (20/40 = 1/2) es mayor que la probabilidad de crecimiento sin fertilizante (20/60 = 1/3).

4. Por tanto, en un análisis probabilístico tanto la luz solar como el fertilizante son causas y la luz solar es el candidato más sólido. Pero, a partir de una descripción de estas posibilidades, como por ejemplo: si hay luz solar, las plantas crecen si se utiliza fertilizante, pero si no hay luz solar entonces las plantas no crecen incluso si se utiliza fertilizante, deberíamos juzgar que la luz solar es la condición que posibilita que el fertilizante cause el crecimiento.

5. La tabla anterior presenta toda la información necesaria para evaluar cualquier probabilidad en este terreno. Por tanto, una teoría probabilística hace desaparecer la distinción de significado entre causas y condiciones posibilitadoras.

▪

* * *

Un mecanismo causal, como muestra el ejemplo de la máquina de coser, no es más que un conjunto de nuevas relaciones causales interpoladas en la cadena que va de la causa al efecto. Podría preguntarse usted por el mecanismo que está detrás de los sencillos efectos físicos que la bobina hace sobre los hilos, y yo podría remitirle a principios mecánicos. Podría usted preguntarse por el mecanismo que está detrás de ellos, etc. Pero no puede haber infinitas causas, lo mismo que la Tierra no puede estar sostenida por una tortuga, que está sostenida por una tortuga, etc., con infinitas tortugas. Por tanto, algunas relaciones causales no tienen detrás mecanismos causales, sino que simplemente suceden. Se sigue que cuando hacemos una afirmación causal, su significado no implica la existencia de un mecanismo causal. Se puede hacer una argumentación similar respecto a los llamados “poderes” causales –la idea de que algunas cosas tienen un poder para provocar efectos–. Como anticipó Hume, nadie ha tenido éxito al formular una definición satisfactoria de poder causal que no presuponga el concepto de causalidad.

Cuando podemos pensar en un mecanismo causal, tiende a ser más importante que los datos sobre las frecuencias de los sucesos, tanto en los experimentos como en la vida real. Esta tendencia puede llevarnos al error. Por ejemplo, se pensaba que los miasmas nocivos eran la causa del cólera y otras enfermedades. El mecanismo era obvio: inspirabas el aire malo y te infectabas por contagio. De manera similar, durante demasiado tiempo se pensó que el rechazo materno era la causa del autismo. Era un mecanismo verosímil –incluso los padres de niños autistas podían imaginar ejemplos en los que habían descuidado a sus hijos–. Estos mecanismos erróneos son convincentes, porque solo se pueden eliminar mediante una cuidadosa investigación científica. El derrocamiento de la teoría de los miasmas es un tema al que vuelvo en el penúltimo capítulo del libro.

* * *

Un misterio práctico que resuelve la presente explicación es el problemático papel de las condiciones posibilitadoras en la ley. La causación está en el núcleo de la justicia porque a aquellos que causan daño a otros les corresponde un castigo o reparar los daños. En el pasado, los teóricos legales habían estado de acuerdo con John Stuart Mill en que la diferencia entre causas y condiciones posibilitadoras es inmoral y caprichosa y por eso con frecuencia en las sentencias legales no se tiene en cuenta o es imprecisa. Consideremos este caso:

El demandado, un constructor, de forma negligente dejó abierto un hueco de ascensor sin vigilancia, y un muchacho sabiendo que el ascensor no estaba se hizo pasar por el encargado del ascensor e invitó al demandante a entrar. ¿Causó el constructor las resultantes heridas del demandante?

La corte decidió que la acción libre de un tercero, el muchacho, bloqueó la conexión causal, y por tanto el constructor no era responsable de los daños.

Un caso opuesto se refiere a los fabricantes de armas de fuego norteamericanos:

Los demandados, los fabricantes de armas de fuego, de forma negligente hicieron unos suministros excesivos a las tiendas de los estados con leyes permisivas respecto a las armas de fuego. Los criminales entraron en posesión de estas armas y asesinaron a familiares de los demandantes. ¿Causaron los homicidios los demandados?

La corte decidió que, al menos algunos de los demandados eran la causa de los homicidios, y por tanto eran responsables de los daños. (La corte sostuvo que eran la “causa más cercana” de los homicidios. La frase se utiliza en la ley consuetudinaria, pero para los estudiosos legales es un misterio qué significa exactamente.) En ambos casos, los demandados posibilitaron que sucedieran los hechos dañinos. La acción negligente del constructor posibilitó el que el chico causara el accidente. Las acciones negligentes de los fabricantes de armas de fuego posibilitaron el que los criminales asesinaran a sus víctimas. Es una pena que la ley de daños en el mundo de habla inglesa no tenga el concepto de que, tanto aquellos cuya negligencia posibilita que se produzca el perjuicio, como los que lo causan, deberían ser responsables de pagar los daños. En lugar de eso, los tribunales fuerzan encontrar una causa donde solo se produce una acción posibilitadora, como en el caso de las armas.

* * *

Si su servicio en el tenis toca la cinta de la red, usted está obligado a servir de nuevo. Es una regla del tenis. ¿Es su obligación moral servir otra vez? Parece extraño decir eso. Pero, si su deseo de ganar en Wimblendon le lleva a hacer trampa, entonces usted ha infringido un código moral, si no las leyes del juego. Si usted va a Wimblendon como espectador entonces usted debe ver un partido en la Pista Central. No es una regla ni una obligación moral, sino una recomendación. Y si usted va a la Pista Central a ver las finales del Campeonato, debe tener una entrada. Es una obligación moral y legal. Puede ser usted demandado en un caso criminal si las autoridades descubren que entra sin ella. Todos estos usos del ”debe” son deónticos. Lo que varía en ellos son los

principios a los que el verbo alude de manera tácita –desde las reglas del tenis a las leyes criminales británicas–. Voy a considerar ahora las bases de la deóntica, y mi intención es mostrar que bajo la misteriosa variedad de temas deónticos hay conceptos sencillos.

* * *

Los lógicos han desarrollado lógicas deónticas basadas en los dos conceptos centrales de obligación y permisividad. Como sus equivalentes en el campo causal –las causas y las condiciones posibilitadoras se basan en los conceptos de necesidad y posibilidad–. Tenemos una necesidad moral de hacer lo que estamos obligados a hacer. Y tenemos una posibilidad moral de hacer lo que nos está permitido hacer. Pero existe una diferencia crucial: en el campo causal lo que es necesario es lo que sucede; pero en el campo deóntico, lo que es obligatorio puede no ser lo que sucede, porque podemos dejar de cumplir nuestras obligaciones. Los dos conceptos deónticos, como los dos conceptos causales, pueden definirse en función uno del otro. Si estamos obligados a salir, entonces no nos es lícito no salir. De la misma manera, si nos está permitido salir, entonces no estamos obligados a no salir.

La teoría de modelos postula que las posibilidades están en el centro del razonamiento, y por tanto una extrapolación sencilla es que las proposiciones deónticas se refieren a posibilidades deónticas, es decir, a estados aceptables. Cada modelo para una proposición deóntica representa, o un estado lícito o, en algunos casos, un estado no lícito. Si una acción es común a todos los modelos de lo que es lícito, entonces es obligatoria. Algunas proposiciones deónticas son categóricas, como el imperativo categórico de Kant: actúa de acuerdo con una máxima que puedas desear que se convierta en una ley general. Pero muchas proposiciones establecen una relación entre estados posibles y lícitos: si su servicio golpea en la cinta de la red usted debe servir de nuevo.

* * *

Los niños empiezan a entender las expresiones deónticas a una edad temprana, y la mayoría empieza a producir ese tipo de expresiones antes de los tres años. Suelen aprender antes lo que no deberían hacer que lo que deberían hacer, y esta observación da lugar a una predicción sobre el razonamiento adulto a que llegaré en el debido momento.

Una proposición como por ejemplo:

Ganar un sueldo te obliga a pagar impuestos.

es compatible con tres estados en los que el primer ítem de cada estado es una posibilidad y el segundo es lo que es lícito:

Estados posibles	*Estados lícitos*
ganar un sueldo	pagar impuestos
no ganar un sueldo	pagar impuestos
no ganar un sueldo	no pagar impuestos

Recordará usted que los modelos tienen siempre algún tipo de estatus epistemológico. Exigen un sistema de etiquetado al que son sensibles nuestros procesos mentales. No voy a seguir etiquetando los modelos en este capítulo, pero debe usted tener presente su estatus. También debe tener presente que en muchos casos el primer estado en un modelo deóntico es ambas cosas: posible y lícito. Lamentablemente, este estatus no es universal, y un ejemplo como: hacer

chanchullos con sus impuestos le obliga a confesar, tiene un estado inicial que es posible pero no lícito.

La teoría postula que los modelos mentales no representan todos los estados posibles y lícitos a los que se refiere una proposición deóntica. Nuestros modelos mentales de una proposición, como por ejemplo: Ganar un sueldo obliga a pagar impuestos, hace más destacada esta situación:

ganar un sueldo pagar impuestos

...

La elipsis denota otras situaciones en las que no se gana un sueldo. La siguiente situación en cuanto a importancia corresponde a lo que no es lícito: ganar un sueldo y no pagar impuestos. Las prohibiciones son similares en cuanto a que una prohibición es una obligación de no hacer algo, pero sus modelos mentales destacan más el caso que no es lícito. Por tanto, una proposición como por ejemplo:

Pagar impuestos impide recibir una pensión

tiene un modelo mental que asocia una posibilidad con aquello que no es lícito (aquí en cursiva):

pagar impuestos recibir una pensión

El siguiente modelo más destacado es el del caso lícito:

pagar impuestos no recibir una pensión

En el caso de una expresión que permite algo, los modelos mentales representan también la acción que convierte en lícita una situación, pero sabemos que también son posibles otras situaciones en las que la acción se produce. Los experimentos han corroborado estas afirmaciones, y el primer estado en que pensaban los participantes correspondía al primer modelo mental explícito para las diversas proposiciones.

* * *

Una vez que las proposiciones se han representado en modelos mentales, se aplican los principios habituales de la teoría de modelos. Podemos extraer conclusiones que corresponden a nuevas relaciones en los modelos resultantes. Consideremos, por ejemplo, cómo podríamos extraer una conclusión a partir de estas premisas:

Ganar un sueldo obliga a pagar impuestos.

Pagar impuestos impide recibir una pensión.

Los principios resumidos en capítulos anteriores (Tabla 8.2) se pueden extender para afrontar los estados lícitos e ilícitos. Dan lugar a un modelo mental más destacado para las dos premisas:

ganar un sueldo pagar impuestos recibir una pensión

en el que la cursiva indica una situación ilícita. Por tanto, las premisas dan lugar a la conclusión:

Ganar un sueldo impide recibir una pensión.

Los modelos completos de las mismas premisas, que le ahorro, dan lugar a la misma conclusión, y por tanto es válida.

Una inferencia muy diferente se basa en premisas de este tipo:

Recibir una pensión impide pagar impuestos.

Pagar impuestos obliga a ganar un sueldo.

El modelo mental inicial de la primera premisa es:

recibir una pensión pagar impuestos

y este modelo se puede actualizar de acuerdo con la segunda premisa como sigue:

recibir una pensión pagar impuestos ganar un sueldo

Este modelo da lugar a la conclusión no válida:

Recibir una pensión impide ganar un sueldo.

Puesto que el modelo de la primera premisa se refiere a lo que está prohibido, mientras que el modelo de la segunda premisa no, podríamos construir un modelo de lo que es lícito a partir de la primera premisa:

recibir una pensión no pagar impuestos

y luego actualizar este modelo de acuerdo con la segunda premisa para dar lugar a dos modelos mentales:

recibir una pensión no pagar impuestos

pagar impuestos ganar un sueldo

Estos modelos no apoyan la conclusión anterior, pero en cambio dan lugar a la conclusión:

Recibir una pensión permite no ganar un sueldo.

Esta conclusión, a diferencia de la anterior, es una inferencia válida a partir de las premisas. Pero las inferencias basadas en modelos múltiples son más difíciles, y pueden superarnos. En un experimento los participantes extrajeron sus propias conclusiones con sus propias palabras, y se enfrentaron a dieciséis problemas diferentes desde el punto de vista lógico. Sus conclusiones, correctas o erróneas, corroboraron la teoría de modelos, que predijo el noventa y uno por ciento de sus respuestas. Los problemas que hacen surgir modelos múltiples, como el ejemplo anterior, deben provocar más respuestas de la forma “no se sigue nada”, porque nos resulta difícil determinar qué conclusión se cumple en los modelos múltiples, si es que hay alguna. Los resultados también corroboraron esta predicción: la respuesta “no se sigue nada” se produjo en el treinta y uno por ciento de los problemas con modelos múltiples pero solo en el dos por ciento de los problemas con un solo modelo.

¿Hay alguna forma de confirmar la afirmación de que utilizamos modelos mentales en el razonamiento deóntico? Consideremos el problema:

Solo se le permite a usted llevar a cabo una de las acciones siguientes:

Acción 1: Coger la manzana o la naranja, o ambas.

Acción 2: Coger la pera o la naranja, o ambas.

¿Se le permite coger la naranja?

Usted está autorizado a realizar o bien la acción 1 o bien la acción 2, pero no ambas. Los modelos mentales de la acción 1 representan lo que es lícito coger:

Manzana

Naranja

Manzana Naranja

Apoyan la conclusión de que está permitido coger la naranja. De la misma forma, los modelos mentales de la acción 2 apoyan la misma conclusión. Por tanto, si utilizamos modelos mentales, entonces deberíamos responder: “Sí, se me permite coger la naranja.” Sin embargo, la respuesta es una ilusión. Si cogiéramos la naranja entonces habríamos llevado a cabo ambas acciones, la 1 y la 2, en contra de las instrucciones según las cuales solo se nos permite realizar una de ellas. A diferencia de los modelos mentales, los modelos completos del problema tienen en cuenta que cuando una acción es lícita la otra no lo es. Estos modelos muestran que son lícitos dos estados: o coge usted solamente la manzana, o bien coge usted solamente la pera. En consecuencia, no puede usted coger la naranja, y por tanto la respuesta correcta al problema es “no”.

Un sencillo problema de control para la ilusión tiene las mismas premisas, pero plantea la cuestión:

¿Se le permite coger la pera?

Los modelos mentales de la acción 2 dan lugar a la respuesta “sí”, y esta respuesta es correcta. Por tanto, quienes razonan deberían acertar incluso sin considerar que si llevan a cabo una acción, no deben llevar a cabo la otra.

Los participantes en un experimento realizaron dos problemas ilusorios y dos

problemas de control. Dieron muchas más respuestas correctas para los problemas de control (noventa por ciento) que para las ilusiones (siete por ciento): 239 de los participantes realizaron mejor los problemas de control que los problemas ilusorios, para ocho no hubo diferencia y solo tres realizaron peor los problemas de control que los problemas ilusorios.

En capítulos anteriores he informado de otros estudios sobre inferencias ilusorias. Ha sido difícil encontrarles un antídoto eficaz, que no sea el enseñar a los participantes a considerar tanto lo que es verdadero como lo que es falso, en el caso de las proposiciones factuales. Sin embargo, el razonamiento deóntico ofrece la posibilidad de un antídoto eficaz. Los problemas formulados en términos de prohibiciones deberían llevar a los participantes a pensar en lo que está prohibido, porque el modelo mental destacado representa ese caso. Este es un problema equivalente al anterior problema ilusorio:

Se le prohíbe a usted llevar a cabo más de una de las siguientes acciones:

Acción 1: Coger la manzana o la naranja, o ambas.

Acción 2: Coger la pera o la naranja, o ambas.

¿Se le permite coger la naranja?

Quienes razonan deberían construir los modelos mentales de lo que es ilícito coger:

Manzana *Pera*

Naranja

Como muestran estos modelos, no es lícito coger la naranja, y por tanto los participantes deberían dar la respuesta correcta, “no”. Se sigue que las prohibiciones deberían reducir las inferencias ilusorias.

El correspondiente problema de control formulado en términos de se prohíbe tiene las premisas anteriores, pero la pregunta es:

¿Se le permite a usted coger la pera?

Los modelos mentales de lo que es ilícito no incluyen la situación de coger solamente la pera, y por tanto quienes razonan deberían dar la respuesta correcta, “sí”. Pero, la tarea debería ser más difícil que un problema de control basado en se permite, que da lugar a un modelo que representa que es lícito coger la pera. Por tanto, la teoría de modelos predice que los problemas ilusorios deberían dar lugar a un mejor rendimiento con se prohíbe que con se permite, pero los problemas de control deberían dar lugar a un mejor rendimiento con se permite que con se prohíbe.

Un experimento corroboró estas predicciones. Los problemas ilusorios dieron lugar a un mejor rendimiento con se prohíbe que con se permite (una mejora del veintiocho por ciento en cuanto a corrección), mientras que para los problemas de control era cierto lo contrario (una disminución del treinta y siete por ciento en cuanto a corrección). Por tanto, aunque el verbo prohíbe palia las ilusiones, lo hace a costa de empeorar la actuación con los problemas de control.

Podría usted sospechar que las inferencias ilusorias llevan a quienes razonan por mal camino porque los problemas son inusuales, forzados, artificiales. Al razonar asumen que se están enfrentando al discurso normal, no a problemas sutiles que son engañosos. Su crítica parece verosímil, pero no resiste un análisis. Por ejemplo, los participantes lo hacen muy bien con los problemas de

control, que se basan en las mismas premisas, y plantean una pregunta similar. La teoría de modelos mentales predice tanto la realización incorrecta de las inferencias ilusorias como la realización correcta de las inferencias de control. También predice que las ilusiones se deberían reducir para problemas de control formulados en términos de prohibiciones.

No se sorprenderá usted mucho al saber que existen teorías alternativas para el razonamiento deóntico. Algunas se basan en reglas formales de inferencia; otras se basan en reglas con un contenido específico o en la invocación de "módulos" innatos en la mente que están especializados en el razonamiento sobre el engaño –la opinión de los psicólogos evolucionistas, que discutí anteriormente (en el Capítulo 18)–. Estas teorías se presentaron para explicar el rendimiento en la tarea de selección –la tarea en la que los participantes tenían que seleccionar pruebas pertinentes para establecer violaciones de principios generales–. Esta tarea ha sesgado las teorías del razonamiento deóntico hasta el punto de que muchas de ellas tienden a no considerar ningún otro tipo de razonamiento. Las observaciones que he descrito en el presente capítulo presentan por tanto un desafío a esas otras teorías. Necesitan explicar los significados de los términos modales el que algunas situaciones lícitas destaquen frente a otras y la aparición de inferencias ilusorias. Este último problema es grave para las teorías basadas en reglas generales. Si las reglas solamente dan lugar a conclusiones válidas, no pueden explicar las inferencias ilusorias; si dan lugar a conclusiones no válidas, corren el riesgo de ser internamente inconsistentes.

* * *

Un sistema deóntico racional, sean las reglas de un juego como el fútbol, la ley consuetudinaria o un sistema ético debería como mínimo basarse en principios consistentes. Suponiendo que es un sistema viable de razonamiento válido, estos principios dan lugar a conclusiones deónticas. Pueden no hacerlo en ciertos casos si los principios no cubren todas las eventualidades, pero no deberían conducir a conflictos respecto a qué es lícito o qué es obligatorio. En realidad, no importa con cuánto ahínco lo intentemos, nos resulta casi imposible diseñar principios deónticos completos y consistentes. Las reglas de muchos juegos son incompletas. Por ejemplo, una vez en el Campeonato Mundial de Fútbol, los

hinchas ingleses se quejaron del tamaño de los guantes del portero alemán. Las reglas del fútbol, lamentablemente, no hablan de los guantes del portero, aunque al menos son consistentes y están libres de paradojas. El sistema legal, no. Un juez inglés, Lord Halsbury, hizo el prudente comentario:

Un caso solo tiene autoridad sobre lo que realmente decide. Niego rotundamente que se pueda citar para una proposición que pueda parecer seguirse lógicamente de él. Esa forma de razonar asume que la ley es necesariamente un código lógico, mientras que todo abogado debe reconocer que la ley no siempre es lógica en absoluto.

En la gran novela de Joseph Heller Trampa-22 el protagonista, el Capitán Yossarian, es un bombardero que vuela en misiones durante la Segunda Guerra Mundial en Italia (como había hecho el propio Heller). Su compañero de tienda, Orr, se vuelve loco en combate, y la locura es motivo para que le liberen de las tareas de combate. Como escribe Heller:

Solo había una paradoja, y era la Trampa-22, que especificaba que el preocuparse por la propia seguridad frente a los peligros que eran reales e inmediatos era el proceso de una mente racional. Orr estaba loco y podía ser retirado del servicio. Todo lo que tenía que hacer era solicitarlo; y tan pronto como lo hiciera, ya no seguiría estando loco y tendría que volar en más misiones. Orr estaría loco para volar en más misiones y cuerdo si no lo hiciera, pero si estuviera cuerdo tendría que volar en ellas. Si volara estaría loco y no tendría que hacerlo, pero si no quisiera hacerlo, estaría cuerdo y tendría que hacerlo. Yossarian se conmovió muy profundamente por la absoluta simplicidad de esta cláusula de la Trampa-22 y soltó un respetuoso silbido.

La ironía final en la novela es que no existe la Trampa-22 pero todo el mundo cree que sí. Esto empeora la situación, porque no hay forma de refutarla.

Algunos aspectos de la ley son análogos a Trampa-22. Por ejemplo, en el caso de

Douglas v. el Estado de California (1963), una norma de California daba a los indigentes derecho a un abogado que les representara en la apelación. Pero este derecho estaba disponible solamente si la Corte de Apelación consideraba que el caso del indigente lo merecía. La corte tenía que decidir si el caso lo merecía para permitir al indigente ser representado mientras la corte decidía si el caso lo merecía. En Dobner v. Peters (un caso planteado en la Corte de Apelación en Nueva York, 1921), la corte determinó que los bebés mutilados cuando eran fetos no daban lugar a acciones legales. Treinta años más tarde, la corte revocó esta decisión. Determinó que Dobner v. Peters ya no seguía siendo una norma correcta, porque su base fundamental había sido la falta de un precedente. Sin embargo, ahora había un precedente, a saber, Dobner v. Peters.

En la parodia de W. E. Bowman sobre una expedición de montañismo británica, The Ascent of Rum Doodle, escaladores intrépidos tratan de subir una famosa montañita conocida como Rum Doodle, que tiene una altura de cuarenta mil pies y medio. Intentan establecer un campamento base para dos escaladores a 39.000 pies. El libro esboza su plan:

El equipamiento para el campo tenía que llevarse desde la cabecera de línea de Chaikhosi, una distancia de quinientas millas. Se necesitarían para ello cinco porteadores. Para transportar la comida de estos cinco se necesitarían dos porteadores, y otro llevaría la comida de estos dos. Su comida la llevaría un muchacho. El muchacho llevaría su propia comida.

Este tipo de razonamiento condujo a los escaladores a inferir que la expedición al completo necesitaría 3.000 porteadores y 375 muchachos. Ese mismo tipo de razonamiento lleva a reglas, como la que determina el impuesto sobre la renta en Estados Unidos:

Si su empleador paga su impuesto sobre la renta, entonces este dinero es una ganancia extra para usted. Por tanto, está sometida a impuesto. Si su empleador paga el impuesto sobre esta cantidad extra, entonces esto también es ganancia y por tanto usted debe pagar impuestos por ella. Si su empleador paga este impuesto, entonces... y así sucesivamente ad infinitum.

Sin embargo, las series se pueden sumar, y la cantidad de impuestos que usted debe resulta ser más de lo que usted pagaría si pagara su propio impuesto de la renta.

* * *

Razonamos para determinar el estatus deóntico de una acción. Puede ser lícita, obligatoria o estar prohibida. También razonamos para determinar si hemos roto una regla o violado un principio moral. Nuestro razonamiento tiene sus límites, y no siempre puede ayudarnos a decidir entre obligaciones que compiten entre sí. Tenemos la obligación de cuidar a nuestros hijos, y tenemos la obligación de ser honrados. Supongamos que no tenemos suficiente dinero para pagar una operación médica que es vital para uno de nuestros hijos, pero entonces encontramos un escondite de dinero ilegal perteneciente a un criminal. Podemos encontrarnos en un dilema sobre si llevarlo a la policía o guardarlo. Tratamos de resolver estos conflictos mediante el razonamiento. Por ejemplo, pensamos interiormente:

Es una cantidad enorme de dinero sucio y mi obligación es llevarlo a la policía. Pero podría devolver todo menos la cantidad que necesito para la operación. ¿Dónde está el daño en eso? El daño está en que está mal, y no debería hacerlo. Sin embargo, el dinero viene de criminales, por tanto, ¿por qué no debería defraudarles una pequeña proporción que nadie más necesita? Porque está mal. Sin embargo, yo tengo una obligación moral de cuidar a mi familia, y esta obligación es más importante que cualquier otra. Si no guardo el dinero, mi hijo sufrirá y puede que incluso muera, y será mi responsabilidad.

El razonamiento nos ha llevado a dos conclusiones incompatibles. Todos experimentamos estos conflictos, y autores que van desde Platón a los creadores de los dibujos animados de Tom y Jerry han representado nuestras luchas con

ellos. Si un sistema de razonamiento válido parte de premisas consistentes, los conflictos son imposibles. Suponer que somos incapaces de razonar de forma válida sería postular el caos en nuestras mentes. Los conflictos sobre lo que debemos hacer se producen por tanto porque nuestro punto de partida no es un único conjunto consistente de creencias deónticas, sino más bien, como en el caso de la causación, varias creencias e intuiciones incoherentes.

Sidney Smith, el agudo clérigo anglicano del siglo dieciocho, caminaba por la ciudad de Londres con un amigo. Llegaron a una calle estrecha en la que los pisos superiores de las casas de ambos lados se inclinaban hacia afuera acercándose uno a otro, y desde dos ventanas enfrentadas dos mujeres estaban gritándose insultos en una discusión sin fin. Smith señaló a su amigo, "Nunca se pondrán de acuerdo –están discutiendo a partir de premisas diferentes–". En una discusión entre Budha y Nietzsche, que Bertrand Russell imaginó, también ellos dos discuten a partir de premisas diferentes. La premisa de Budha era que cualquiera que sufre le concierne; la premisa de Nietzsche era que solo el sufrimiento de ciertos individuos le concierne. Debido a que sus premisas eran diferentes, como señaló Russell, su conflicto no se podría resolver de una forma racional.

La mayoría de nosotros estamos hechos con un poco de Budha y un poco de Nietzsche. Nuestras discusiones interiores se producen a partir de principios que afectan a todo el mundo y principios que afectan a ciertas personas especiales en nuestras vidas. Y por tanto el razonamiento tampoco puede resolver siempre nuestros conflictos morales. Puede que eso no sea malo. Algunos de los individuos más aterradores son quienes han conducido sus vidas guiados por dogmas únicos inflexibles: Torquemada, Robespierre, Hitler, Pol Pot, Bin Laden.

La mayoría de nosotros tenemos intuiciones morales inmediatas derivadas de inferencias inconscientes del tipo de las que he descrito en la Parte II de este libro. Estas inferencias son computaciones más rudimentarias que las que corresponden a las inferencias conscientes. Las propias intuiciones puede que no estén libres de conflicto, y nos pueden permitir esquivar nuestra responsabilidad o desconectar nuestro sentido de la moralidad de acciones desagradables que tenemos la obligación de realizar. Sin embargo, cuando nuestras intuiciones son homogéneas, pueden hacer que cambiemos nuestras decisiones razonadas. Ciertamente, algunos autores las consideran inspiradoras de nuestros juicios morales.

Los psicólogos, incluidos algunos de mis colegas de Princeton, han examinado nuestras intuiciones morales. Han comparado dilemas en los que podemos actuar en la forma impersonal de un bombardero que lanza bombas con dilemas en los que tenemos que actuar como un soldado que utiliza un arma blanca en una batalla sangrienta. Un ejemplo de este tipo de dilema es:

Un tranvía fuera de control está a punto de atropellar y matar a cinco personas. Usted puede accionar un interruptor que llevará el tranvía a un conjunto de vías diferente donde matará a una persona en lugar de a cinco. ¿Accionaría usted el interruptor?

La mayoría de la gente dice, “sí”. Pero supongamos que la única forma de salvar a las cinco personas es que usted empuje a una persona delante del tranvía, de forma que muera en lugar de las cinco. ¿Lo haría usted? La mayoría de la gente dice, “no”. Esta diferencia puede surgir porque nuestro razonamiento moral consciente puede seguir la máxima utilitaria de “el mayor beneficio para la mayoría”. Pero la idea de tener que empujar a alguien a la muerte hace surgir una emoción, y esta afecta a nuestra decisión. Esta diferencia se puso de manifiesto en un estudio sobre imágenes cerebrales: se activaban diferentes zonas en el cerebro de los participantes al tomar ambos tipos de decisión. Ciertamente, nuestras emociones pueden abrumar a nuestras decisiones razonadas o incluso adelantarse a ellas, porque pueden suscitar una respuesta antes de que hayamos acabado de razonar.

Mi anterior ejemplo de conflicto entre la honradez y el dinero para la operación de un hijo surge porque nuestras intuiciones morales o nuestros principios conscientes son inconsistentes. Ambas opciones tienen sus bases emocionales. Resolvamos como resolvamos el conflicto, podemos inferir que actuaremos mal. En un caso, no hemos sido honrados y eso está mal; en el otro caso, hemos sacrificado a un hijo por nuestros principios y eso también está mal. De cualquier forma, sentimos que lo hemos hecho mal.

* * *

La maquinaria para razonar sobre la causación y la obligación parece ser la misma que la maquinaria para razonar sobre cualquier otro asunto. Construimos modelos, a partir de ellos obtenemos conclusiones y, quizá, buscamos contraejemplos para ellas –especialmente si otros nos proporcionan conclusiones para evaluar–. Por consiguiente, no parecen existir procesos especiales de razonamiento causal o deóntico. Es el mismo asunto de siempre. Sin embargo, algunos aspectos de estos tipos de razonamiento parecen especiales. Las asunciones a partir de las que razonamos están plagadas de conflictos, que pueden ser más frecuentes que en otros campos del razonamiento. Y las decisiones morales parecen tener más tendencia a suscitar emociones. Lo que es peor, no tenemos una forma sistemática de resolver nuestros dilemas morales. El conflicto y la emoción hacen al razonamiento deóntico diferente del razonamiento sobre los Sudokus.

CAPÍTULO 23

Creencias, herejías y cambios de opinión

De niño, sir Edmund Gosse, el escritor eduardiano, fue educado como miembro de una estricta secta cristiana. Su padre fue el autor de Omphalos, una crítica a la teoría de la evolución, que argumentaba que cuando Dios creó el mundo, incluyó los fósiles, los estratos geológicos, los anillos de los árboles y muchos otros elementos como el ombligo de Adán (el título del libro es ombligo en griego), que no eran un verdadero registro del origen de la tierra. Su padre le dijo al joven Edmund que el peor pecado de todos era la idolatría. Un día, cuando Edmund estaba solo en casa, puso una silla sobre una mesa, inclinó la cabeza y declaró, "Yo te adoro, oh silla". Esperaba ser abatido por un rayo debido a su impiedad. Miró por la ventana. El sol seguía brillando en el claro cielo azul. Y no ocurrió nada.

¿Qué conclusión supone usted que sacó Edmund de esta experiencia?

Su fe religiosa no se vio afectada. Pero, como escribió más tarde, llegó a la conclusión de que su padre podría no tener razón en todo.

La líder de una secta americana del Juicio Final anunció que el mundo estaba llegando al final. En el día señalado, muchos de los fieles se reunieron con ella para rezar en la cima de una montaña. La hora terrible se acercó; pasó. Y no sucedió nada.

¿Qué conclusión supone usted que sacaron los fieles a partir de esta experiencia?

La líder anunció que el mundo se había salvado gracias a sus oraciones y se convirtieron en unos creyentes aún más devotos, e hicieron proselitismo a favor del movimiento aún con más fuerza. Solo los que no habían estado en la cima de la montaña mostraron que su convicción se debilitara.

Usted y yo vemos a un pollo cruzar la carretera. Y por tanto ambos creemos que el pollo cruzó la carretera. Sabemos que nuestras percepciones pueden ser

falibles, pero el hecho de que los dos fuéramos testigos del mismo hecho corrobora nuestra mutua creencia. Sin embargo, una fuente de creencias más potente es, no lo que percibimos, sino lo que otros nos cuentan. No sacrificamos nuestras vidas por una creencia sobre algo que vimos. Pero podríamos hacer ese sacrificio por una creencia basada en el testimonio de otros. El testimonio vital se refiere a la fe como la describía San Pablo: “la sustancia de las cosas que se esperan, las pruebas de cosas no vistas”. Las cosas no vistas, por definición, no se pueden invalidar mediante la observación. Por tanto, algunos cristianos primitivos se alegraban de lo imposible de sus creencias. Creemos, porque otros creen. O, como señaló William James: “Nuestra fe es fe en la fe de otro.” Si queremos crear una religión, deberíamos creer el testimonio que damos y nuestros discípulos deberían estar predispuestos a creernos. Nosotros deberíamos ser creíbles: ellos deberían ser crédulos. Podemos ser todo lo convincentes que queramos tratando de persuadirles de que existe una estrella hecha de cartón, pero, ¿a quién le importa? La fe se refiere a asuntos importantes. Necesitamos proporcionar una visión del mundo que dé una razón para vivir, consuelo para las pérdidas, y esperanzas para una vida futura. San Ignacio de Loyola, el fundador de la orden de los Jesuitas, dijo: “Dadme un niño hasta que tenga siete años y es mío para toda la vida.” Por tanto, deberíamos empezar con niños pequeños. Deberíamos rodear a nuestros novicios de otros creyentes convencidos para proporcionarles apoyo emocional y presión para seguir las directrices. Deberíamos aislar a los novicios de más edad de los no creyentes. Deberíamos darles un nuevo nombre, una nueva identidad, una nueva fe. Esta nueva fe debería incluir ritos simbólicos con una significación emocional. Debería también amenazar a los reincidentes con un terrible castigo.

Deberíamos promulgar un conjunto de creencias que fueran consistentes, autosuficientes e imposibles de refutar. Como ejemplo, consideremos los principios centrales de las sectas terroristas, sean musulmanas, judías o cristianas:

• Nuestro propósito es lograr un mundo perfecto.

• Nuestra tierra es sagrada.

• Debemos protegerla.

• Nuestros enemigos la amenazan.

• Debemos matarles.

• Son seguidores de Satán.

• Sus muertes purifican el mundo.

• Nuestra causa requiere nuestro sacrificio.

• Nuestro sacrificio es nuestro más alto objetivo.

• Obtendremos nuestra recompensa en el cielo.

La difusión de una visión del mundo comienza en una secta. Una secta que tiene éxito se convierte en una religión. Una secta que no tiene éxito se convierte en una herejía. Su fracaso rara vez depende de un fracaso de la fe, sino del poder de la ortodoxia. Los heterodoxos darán sus vidas por la herejía. Para los que están fuera de la batalla de las creencias, las diferencias no parecen ser un asunto de vida o muerte. Consideremos, por ejemplo, estos diferentes artículos de fe:

Jesús era dos personas, una divina y una humana.

Jesús tenía dos naturalezas, una humana y una divina.

Jesús tenía solo una naturaleza, y era divina.

Jesús tenía solo una naturaleza, y era un agente humano de Dios.

Hay muchas otras opiniones sobre la naturaleza de Cristo, pero, ¿cuál de las anteriores creencias es ortodoxa y cuáles de ellas son herejías? Si usted es cristiano, debería saberlo, pero puede comprobar la respuesta en las notas al capítulo. Si usted no es cristiano, podría sorprenderse de cómo podría alguien establecer la opinión correcta. La respuesta es el razonamiento humano.

Las herejías demuestran el poder de las creencias, porque los adeptos tenaces mueren por su doctrina. En la Edad Media, la mayor amenaza para la iglesia católica fue la herejía albigense. Sus adeptos vivían en el Languedoc en lo que ahora es el rincón suroeste de Francia, y se les conocía como Cátaros. Tenían la creencia gnóstica de que el mundo tangible era la obra del diablo, y por tanto Jesús solo aparentaba tener un cuerpo humano. El verdadero Dios era invisible. Los Cátaros trataban de hacer de sí mismos seres morales perfectos. Si no lo conseguían, deberían reencarnarse de nuevo una y otra vez hasta lograrlo. Existían algunos “perfectos” y predicaban el camino de la santidad.

La herejía amenazaba al Catolicismo, porque rechazaba las proclamas espirituales y temporales de la iglesia. El Papa Inocencio III declaró una guerra santa contra los Cátaros. Fue la primera Cruzada. Y en 1209 un ejército normando se extendió por el Languedoc para ocuparse de los Cátaros. El ejército sitió la ciudad de Béziers, un centro de la herejía, y pidió que los ciudadanos entregaran a los perfectos. Los ciudadanos se negaron a hacerlo. Cuando se rompió el sitio, el jefe del ejército normando preguntó al monje cisterciense que estaba a cargo de la cruzada qué hacer. Se dice que respondió, “Matarlos a todos, Dios conocerá a los suyos”. En una mañana de trabajo, los Cruzados degollaron de quince a veinte mil hombres, mujeres y niños. Ambos bandos en lucha estaban al servicio de sus creencias religiosas. El Papa creó luego la Inquisición para extirpar a los Cátaros. Tuvo éxito, pero costó más de un siglo hacer desaparecer la herejía.

¿Por qué estas creencias imposibles de comprobar son tan poderosas como para que aquellos que las apoyan mueran antes de abandonarlas? La respuesta, sugiero, se basa en una transición inconsciente a una emoción. Las creencias se refieren a cómo deberíamos vivir, la naturaleza de la muerte y la supervivencia después de ella. Estos temas tienen una extraordinaria capacidad para invocar una emoción básica de apego. Nos apegamos al padre Dios y a la madre iglesia. Y las emociones básicas, como sugerí en capítulos anteriores, se crean mediante evaluaciones sencillas que llevan a transiciones inconscientes. Cualquier reto a estas creencias primigenias es una profunda amenaza, y nos resistiremos a él.

Experimentamos esta consecuencia si criticamos cualquier valor no comprobable, sea una doctrina religiosa, una opinión política como el marxismo o el capitalismo del laissez-faire, una proposición moral sobre la eutanasia o la vivisección o incluso una teoría científica que se ha convertido en una cuasireligión –como señaló Oscar Wilde, la ciencia es la historia de las religiones muertas–. Quienes la proponen tienen una reacción emocional. La amenaza puede dar lugar a una transición inconsciente a una intensa angustia, y deja de ser posible un intercambio racional de ideas.

Por tanto, ¿qué nos lleva a cambiar nuestras creencias? Saulo en el camino de Damasco –con la intención de perseguir a los cristianos– fue rodeado por una brillante luz, y oyó la voz de Jesús reprendiéndole. Estuvo tres días ciego, y sin comida ni bebida. Y luego "se le cayó la venda de los ojos", y se convirtió en Pablo, el apóstol cristiano. Las creencias sobre cosas que no se ven son excepcionales.

* * *

Algunas creencias respecto a las que tenemos un lazo emocional son evaluables. Y algunas veces persistimos en ellas a pesar de las pruebas. El psicólogo desapasionado dice que tenemos un estereotipo. Supongamos, por ejemplo, que creo que las mujeres son malas conductoras. Puede que no me libre de mi convicción si usted me señala que las compañías aseguradoras cobran primas menores a las mujeres conductoras que a los hombres, porque las mujeres tienen menos accidentes. Las mujeres, podría argumentar, no conducen tantas millas como los hombres, y por tanto participan con menos frecuencia en accidentes. Los estereotipos sobre los sexos, las razas y los grupos sociales persisten ante pruebas conflictivas. Y algunos estereotipos tienen una lógica que les hace difíciles de refutar. Observo un ejemplo de mala conducción, y miro para ver el sexo del conductor. Si el conductor es una mujer, digo, ¡típico! Y mi prejuicio se refuerza. Si el conductor es un hombre, digo, "Bueno, también algunos hombres son malos conductores". Por tanto, ¿qué podría observar que pudiera librarme de mi estereotipo? No es obvio. Pero, en realidad, tengo que ver a mujeres conduciendo bien. El conducir bien, sin embargo, no es tan evidente como el conducir mal. La mejor prueba de la calidad de la conducción a largo plazo es

actuarial, estamos otra vez en las primas de los seguros. Carecen de convicción para cambiar a un escéptico recalcitrante.

Nuestra cultura nos proporciona normas, y puede resultarnos difícil aceptar que en otro sitio hagan las cosas de otra forma. El sesgo es difícil de estudiar en el laboratorio, pero está corroborado por muchas observaciones antropológicas. Una distancia pequeña en la geografía puede ser un gran viaje en la cultura. En Italia, un documento llamado una “residenza” es un registro oficial de dónde vive uno. Le capacita para votar, para tener médico en el sistema nacional de salud, para matricular el coche y para hacer muchas más cosas. El autor Tim Parks trató de explicar a un policía italiano que en el Reino Unido no existe ningún documento de este tipo. Mantuvieron este diálogo, que empezó con el policía haciéndole la pregunta:

¿Así que qué hace uno cuando se traslada?

Se traslada.

¿Y las placas de matrícula del coche?

Las deja como están.

¿Y su carnet de identidad?

No existen los carnets de identidad.

¿Y el médico?

Uno va y se inscribe en la consulta del médico más cercano.

No creyó a Parks: trasladarse no podía ser tan fácil.

Otro mecanismo que protege nuestras creencias de la revisión es la forma en que razonamos. Pregúntese, por ejemplo, qué se sigue de estas premisas:

Todos los franceses que están en el restaurante son gourmets.

Algunos gourmets beben vino.

La mayoría de la gente extrajo la conclusión:

Algunos de los franceses que están en el restaurante beben vino.

Es una conclusión verosímil, porque encaja en nuestro estereotipo sobre los franceses. Pero, pregúntese ahora qué se sigue de estas premisas:

Todos los franceses que están en el restaurante son gourmets.

Algunos gourmets son italianos.

Es improbable que infiera que algunos franceses son italianos.

Los experimentos han corroborado esta diferencia, y la teoría de modelos la predice. En el primer caso, construimos un modelo de las premisas que da lugar a la conclusión: algunos franceses beben vino. La conclusión es creíble, y por tanto hacemos la inferencia sin darnos cuenta de que no es válida. En el segundo caso, construimos un modelo similar y da lugar a la conclusión de que algunos franceses son italianos. La conclusión es absurda, a pesar de la existencia de la Unión Europea, y por tanto buscamos con más insistencia un contraejemplo en el que no se cumpla. Encontramos uno: los gourmets que son franceses pueden ser diferentes de los gourmets que son italianos. Este modelo refuta nuestra conclusión, y por tanto no la extraemos. Por supuesto, un ejemplo similar rebate nuestra conclusión en el primer ejemplo: los gourmets que son franceses son distintos de los gourmets que beben vino. Pero nunca buscamos este contraejemplo, porque estábamos satisfechos con nuestra conclusión inicial. En general, buscamos con más insistencia contraejemplos para conclusiones increíbles que para conclusiones creíbles. Por tanto, es probable que nuestro método de razonamiento permita que nuestras creencias firmes no se vean afectadas por los contraejemplos incluso en aquellos casos en que estos existen, aunque nuestras creencias también pueden empezar por dificultarnos la construcción de un modelo.

* * *

Todos albergamos creencias que resisten la revisión; pero también tenemos muchas creencias que son fáciles de cambiar. Consideremos este típico episodio:

Estoy sentado con María en el exterior de un agradable café en la Provenza en el sur de Francia. Sabemos que Paolo y Vitorio fueron a buscar el coche para recogernos. Sabemos también que si fueron a buscar el coche entonces deberían volver pronto, porque lo dejamos a poca distancia subiendo una colina. Pasan diez minutos, y luego otros diez, sin señales del coche.

Ha sucedido algo inesperado: la vida, lamentablemente, ha conducido a un conflicto con una consecuencia válida de nuestras creencias. No se puede dudar de la realidad. Inferimos que volverían pronto, y no ha sido así. Por tanto, hay que cambiar algo. Como mínimo, nuestra conclusión es falsa, y debemos retirarla. Era una consecuencia válida de nuestras creencias, por tanto, al menos una de ellas también debe ser falsa. O bien no fueron a buscar el coche, o si lo hicieron, no les lleva unos minutos el volver.

Un asunto más urgente es qué ha ido mal y qué deberíamos hacer. ¿Deberíamos ir a buscarles o permanecer donde estamos? El coche es un viejo y cómodo Citroen. Sospechamos que Paolo y Vittorio pueden haber tenido problemas para arrancarlo, porque está dando muestras de su edad. Por tanto, ¿deberíamos subir la colina para ayudarles? No, decidimos, deberíamos quedarnos donde estamos, porque si consiguen arrancar el coche, vendrán al café a recogernos y pueden venir por otro camino. Haríamos mejor en esperarles fuera del café. Y si hay algún otro problema –quizá un sistema complicado de calles de una dirección que les ha obligado a salir de la ciudad–, también haríamos mejor en esperarles. Por tanto, nuestro razonamiento, basado en unos cuantos elementos de razonamiento general, nos lleva a quedarnos donde estamos, admirando la vista del Monte Sainte-Victoire y el castillo donde vivía Picasso. Efectivamente, después de otros diez minutos, el coche aparece resoplando ante nuestra vista. Habían necesitado que les remolcaran para conseguir arrancarlo.

El ejemplo es típico. Se produce una inconsistencia entre nuestras creencias y un hecho incontrovertible, y estamos obligados a cambiar nuestras creencias. El mismo tipo de inconsistencias se producen en la ciencia. Creemos, pongamos por caso, que el calor es una sustancia y que las sustancias tienen peso. Calentamos un objeto, pero descubrimos que no es más pesado. Necesitamos explicar esta inconsistencia, y la solución exige razonamiento. Tenemos que revisar al menos una de nuestras creencias. También tenemos que intentar razonar la forma de volver a la consistencia. Necesitamos hacer una abducción que resuelva la inconsistencia. Nuestro razonamiento no es un mero ejercicio académico. Afecta a lo que decidimos hacer. Estos tres pasos –la detección de una inconsistencia, la revisión de las creencias y la explicación de la inconsistencia– proporcionan al psicólogo un orden del día para el estudio de los cambios de opinión.

* * *

Cuando leemos una descripción de algunos hechos, construimos un modelo de lo que sucedió. Si en algún punto la descripción es inconsistente con nuestro modelo de la situación, podemos darnos cuenta de la discrepancia. La novela de Daniel Defoe Robinson Crusoe describe que el héroe del mismo nombre se arranca la ropa y nada hasta el barco naufragado para rescatar cosas de valor. No mucho después de llegar a bordo, Defoe le tiene "llenándose los bolsillos de galletas". Pero, si Crusoe está desnudo, no tiene bolsillos. El crítico literario John Sutherland se dio cuenta de esta inconsistencia, y ha recopilado varios libros sobre errores cómicos que aparecen en novelas famosas. Una lectura detenida de la Biblia también revela inconsistencias evidentes. El Génesis, por ejemplo, implica tanto que Dios creó los árboles antes que al hombre como que Dios creó los árboles después que al hombre. Frank Kermode, el decano de los críticos literarios, ha señalado otros errores triviales que han cometido los novelistas. Tolstoy confundió las fechas en Guerra y Paz. Dickens, a pesar de mantener cuidadosos memorándums sobre sus personajes, se olvidó del perro Diógenes en la parte final de Dombey e hijo. Como señala Kermode, en la gran literatura se producen grandes inconsistencias, aunque a veces se utilizan como efecto literario. Necesitamos una vista de águila y una memoria de elefante para detectarlas. Quizá Sutherland adoptó una actitud crítica con la que las buscó de forma deliberada –espero que lo hiciera, porque sería horroroso no ser capaz de leer de otra forma–. Kermode habla de la "atención anormalmente cuidadosa" con la que leen los especialistas en literatura.

Nos ocupamos de la consistencia con un éxito variable. La vida siempre nos está sorprendiendo, y una sorpresa es algo inconsistente con nuestras creencias o sus consecuencias. Y puede ser vital cambiar de opinión a la luz de una sorpresa. Si nuestro razonamiento ha sido impecable –y puede no haberlo sido–, entonces al menos una de nuestras creencias es falsa. La lógica, lamentablemente, no nos dice cuál es la creencia infractora.

* * *

Decir que una serie de proposiciones es consistente significa que es posible que todas ellas sean verdaderas. Como corolario, una serie de proposiciones es inconsistente si es imposible que todas ellas sean verdaderas. Un caso especial de inconsistencia es una auto- contradicción. Yogi Berra, el famoso jugador de béisbol yanqui, se especializa en expresiones de este tipo, como por ejemplo, “realmente no dije todo lo que dije”. Si una serie de proposiciones es inconsistente, entonces al menos una proposición de la serie es falsa. Y si actuamos en base a una creencia falsa, puede producirse un desastre.

En lógica, la inconsistencia tiene consecuencias drásticas: podemos probar cualquier proposición partiendo de una inconsistencia. Bertrand Russell en la presidencia de una cena en el Trinity College, Cambridge, encontró a un entrevistador escéptico, que le preguntó si esta consecuencia de las contradicciones era cierta. Russell replicó que lo era. En ese caso, el escéptico dijo, pruebe que yo soy el Papa a partir de 1 + 1 = 1. Y Russell dijo: Usted es uno. El Papa es uno. Uno más uno igual a uno, y por tanto usted y el Papa son uno. Una ingeniosa réplica, aunque la verdadera prueba de que las inconsistencias implican cualquier proposición es un poco más técnica.

* * *

Hasta este momento, puede que usted haya asumido que la inconsistencia se produce solo cuando una proposición entra en conflicto con otra. Pero también se puede producir entre una serie de proposiciones en la que no hay dos proposiciones que entren en conflicto. He aquí un ejemplo sencillo:

Si Alan está en mi fiesta entonces también está Betty.

Si Betty está en mi fiesta entonces también está Chuck.

Alan está en mi fiesta pero Chuck no.

Con pensarlo un momento debería usted convencerse de que las tres proposiciones no pueden ser verdaderas, pero cualquier pareja que formemos puede serlo. Podemos rehacer este ejemplo con mil proposiciones o con un millón, con la condición de que en el momento en que eliminemos una única proposición, entonces las restantes 999 o 999.999 proposiciones, sean las que sean, formen una serie consistente.

Las proposiciones que se afirman en este libro son inconsistentes. Pensé que era lo mejor decirlo para anticiparme a mis críticos. La afirmación debe ser cierta, puesto que si todas las demás proposiciones del libro son consistentes, entonces la propia afirmación es inconsistente con ellas, y por lo tanto verdadera. Y si todas las otras proposiciones son inconsistentes en sí mismas, entonces la proposición también es cierta. El libro contiene miles de proposiciones sencillas, como por ejemplo, "Tostoy confundió las fechas en Guerra y Paz", y asegura varias relaciones entre ellas. Es consistente si todas esas relaciones son verdaderas al menos en una posibilidad. Un centenar de proposiciones sencillas da lugar a tropecientas posibilidades, porque cada proposición puede ser o bien verdadera o bien falsa. En el peor de los casos, podríamos tener que buscar entre todas esas posibilidades. El número parece inofensivo, pero no lo es. Es tan grande que incluso si pudiéramos comprobar cada posibilidad en una millonésima de segundo, nos seguiría llevando más de cuarenta mil millones de millones de años comprobarlas todas. Por tanto, las pruebas de consistencia son inabarcables. Exigen cantidades más y más grandes de tiempo y de memoria a medida que el número de proposiciones sencillas diferentes aumenta. ¡Buena suerte con la comprobación de que las proposiciones de este libro son consistentes!

La inabarcabilidad de un campo no significa que cada uno de los problemas que se incluyen en él sea imposible de resolver. Los problemas a pequeña escala se pueden resolver tanto mentalmente como con ordenadores. Una forma obvia de economizar es que nuestras creencias tienden a estar separadas en grupos independientes. Nuestras creencias sobre Jesús, pongamos por caso, son independientes de nuestras creencias sobre las normas del tenis. (Ahora, por supuesto, usted sí tiene una creencia sobre ambas –una creencia de segundo orden respecto a que usted no tiene ninguna creencia de primer orden sobre

ambas–). Esta segregación nos permite también mantener visiones del mundo aisladas pero inconsistentes. Las distintas culturas parecen mantener ese tipo de opiniones, y nosotros también. He encontrado a un psicólogo experimental que creía en la curación por la fe, a un ingeniero que creía en la ESP (percepción extrasensorial) y a un físico que creía en cucharas que se doblan. Y ¿quién sabe qué extrañas acumulaciones encontraríamos en su mente o en la mía? Quizá el poeta Walt Whitman se estaba regocijando con esa variopinta diversidad cuando escribió: “¿Me contradigo a mí mismo? Muy bien, me contradigo a mí mismo...”

* * *

Nos damos cuenta de las inconsistencias flagrantes, pero, ¿cómo detectamos las más sutiles? Las pruebas sugieren que utilizamos modelos mentales. Buscamos un modelo de una posibilidad en la que todas las proposiciones son verdaderas. Si encontramos ese modelo, evaluamos las proposiciones como consistentes; si no conseguimos encontrar ese modelo, evaluamos las proposiciones como inconsistentes. Dado que podemos enfrentarnos como máximo con solo un puñado de proposiciones, una estrategia es empezar con un modelo para una posibilidad que satisfaga la primera proposición de una serie, y luego comprobarlo frente a cada una de las sucesivas proposiciones. Si encontramos una proposición que excluye la posibilidad en la que estamos pensando, entonces empezamos de nuevo. La Tabla 23.1 ilustra la descripción de cómo llevamos a cabo la tarea según la teoría de modelos, utilizando un procedimiento que he implementado también en un programa de ordenador.

Tabla 23.1. Pruebas de consistencia utilizando modelos mentales de las posibilidades

1. La tarea es decidir si un conjunto de frases es consistente o no. Dado el siguiente conjunto de oraciones que describen la fruta que hay en un mostrador:

Si no hay una manzana entonces hay un plátano.

Si hay un plátano entonces hay una cereza.

No hay una manzana y hay una cereza.

construimos un modelo mental de la posibilidad más destacada para la primera oración:

no manzana plátano

Este modelo es compatible con la segunda oración, que lo actualiza:

no manzana plátano cereza

Este modelo es compatible con la tercera oración, y por tanto juzgamos que las oraciones son consistentes.

2. Dado un conjunto de oraciones, equivalente desde el punto de vista lógico al primer ejemplo:

Hay una manzana o hay un plátano, o ambas cosas.

No hay un plátano o hay una cereza, o ambas cosas.

No hay una manzana y hay una cereza.

construimos un modelo mental de la posibilidad más destacada para la primera oración:

manzana

La ausencia de un plátano en este modelo es consistente con la segunda oración, que actualiza el modelo:

manzana no plátano

Este modelo no es consistente con la tercera oración, y por tanto tenemos que volver atrás y construir un nuevo modelo para la primera oración:

plátano

que la segunda premisa actualiza:

plátano cereza

Este modelo es compatible con la tercera premisa, que lo actualiza:

no manzana plátano cereza

La necesidad de buscar una posibilidad alternativa predice que este problema debería ser más difícil que el primero.

Las pruebas apoyan la explicación que aparece en la tabla. Los participantes en un experimento tenían que decidir si era posible que todas las oraciones de una serie fueran verdaderas –decidir, en realidad, si las oraciones eran consistentes–. Eran más exactos con series de condicionales que con series de disyunciones, similares en ambos casos a las que aparecen en la tabla. Estos resultados sugieren que probaban los modelos de uno en uno, y tenían que volver atrás para buscar un modelo alternativo para los problemas disyuntivos. Cuando las oraciones eran inconsistentes, la tarea era un poco más difícil, como también predice la teoría de modelos, pero la diferencia entre los dos tipos de problemas se reducía. Otro estudio corroboró la utilización de modelos mentales mostrando que los participantes sucumbían ante las inferencias ilusorias. Los problemas tenían una instrucción inicial que dejaba claro que una oración era verdadera y una oración era falsa, p. ej.:

Solo una de las proposiciones siguientes es verdadera:

La bandeja es pesada o elegante, o ambas cosas.

La bandeja es elegante y portátil.

La proposición siguiente es verdadera sin ninguna duda:

La bandeja es elegante y portátil.

Escriba una descripción de la bandeja_____

Los participantes escribían una descripción si evaluaban las premisas como consistentes. Como predice el modelo mental de las premisas, tendían a juzgar que en este caso las premisas eran consistentes y a describir la bandeja como: ligera, elegante y portátil. En realidad, las premisas eran inconsistentes. Si la bandeja fuera elegante y portátil, ambas oraciones de la disyunción inicial serían verdaderas, en contra de la instrucción según la cual solo es verdadera una de ellas. En el experimento, los participantes sucumbieron a las ilusiones, pero realizaron mejor los problemas de control. Generaron descripciones para las ilusiones y los respectivos problemas de control que correspondían a los modelos mentales de las proposiciones.

* * *

¿Qué deberíamos hacer cuando detectamos una inconsistencia? En lógica, un hecho sorprendente es que si de unas premisas se sigue una conclusión, entonces ninguna premisa adicional puede nunca anular la validez de esta inferencia. En

mi anterior ejemplo, María y yo hicimos la deducción válida:

Paolo y Vittorio han ido a recoger el coche.

Si Paolo y Vittorio han ido a recoger el coche entonces volverán pronto.

Por tanto, Paolo y Vittorio volverán pronto.

¿Qué premisa adicional podría invalidar esta inferencia? Podría usted pensar: bueno, el hecho de que Paolo y Vittorio no vuelven pronto. Pero no, el añadir esta premisa da lugar a una serie de premisas contradictorias, y en lógica a partir de premisas contradictorias se sigue cualquier conclusión, sea la que sea. Y por tanto, por supuesto, también se sigue de ellas la conclusión anterior. Por consiguiente, a medida que añadimos nuevas premisas a una serie ya existente, lo único que ocurre es que podemos sacar nuevas conclusiones. Los lógicos utilizan un término de su jerga para este hecho. Dicen: la lógica es monotónica, porque con cada premisa adicional se siguen más conclusiones a partir de las premisas. La lógica, no el amor, significa no tener que decir nunca lo siento. Nunca tenemos que retractarnos de nada.

En la vida real, estamos inclinados a retirar una conclusión si choca con el muro de un hecho inconsistente. Para utilizar otro término artístico, nuestro razonamiento es no monotónico. Retiramos una conclusión previa, incluso si su inferencia es válida, a la luz de un posterior conflicto con la realidad. La retirada de conclusiones está omnipresente en la vida, debido a la frecuencia con que los sucesos conspiran para derrotar a nuestras inferencias. Pero, ¿cuál de nuestras creencias abandonamos?

En algunos casos, podemos eludir esta cuestión. A menudo razonamos a partir de idealizaciones, como: un perro tiene cuatro patas. Consideramos que nuestra idealización es verdadera a menos que existan hechos en contra. Si existe un hecho que esté en contra – mi perro Figo, por ejemplo, solo tiene tres patas–,

abandonamos las consecuencias en este caso, pero mantenemos la idealización.

La inferencia sobre Paolo, Vittorio y el coche no dependía de ninguna asunción por defecto. María y yo no teníamos más opción que revisar nuestras creencias, y la lógica no nos puede decir qué premisa deberíamos retirar. Una de nuestras premisas era condicional: si Paolo y Vittorio han ido a recoger el coche, entonces volverán pronto. La otra era categórica: Paolo y Vittorio han ido a recoger el coche. Si combinamos nuestra premisa condicional con el hecho de que no vuelven pronto, entonces se deduce que no fueron a recoger el coche. Podríamos entonces retractarnos de nuestra premisa categórica. Otra posibilidad, sin embargo, es que la premisa condicional en sí misma sea falsa.

Las creencias se diferencian en lo arraigadas que están. Ya he hablado de las inversiones emocionales en creencias, y por tanto uso la palabra “arraigo” para referirme únicamente a las bases intelectuales de una creencia, tales como su coherencia con otras creencias, la existencia de razones que la apoyan, su probabilidad, etc. Si una creencia está menos arraigada que otras, quizá deberíamos abandonarla como resultado de un conflicto. Podemos detectar este tipo de creencia débil. Supongamos, por ejemplo, que María y yo hubiéramos sabido que Paolo y Vittorio eran unos ávidos coleccionistas de mariposas, y que siempre salían corriendo a capturar nuevas mariposas que añadir a sus colecciones. Cuando no volvieron pronto, podríamos haber pensado que habían visto una mariposa atractiva y habían ido a capturarla. Por tanto, abandonamos nuestra creencia de que fueron a recoger el coche. Los psicólogos han comprobado ideas análogas a esta en experimentos y han corroborado que se producen.

Con el conocimiento general se produce un efecto similar. En general, si nos enfrentamos a la opción de abandonar un principio general o abandonar un hecho concreto, pero supuesto, abandonamos el hecho supuesto. Sabemos que el agua se congela para formar hielo. Tenemos muchas razones para creer en esta proposición. Encaja con nuestras otras creencias sobre el invierno, los frigoríficos y el montañismo, y tenemos muchas observaciones que la apoyan. Por tanto, si alguien nos dice: congelé agua, pero no se convirtió en hielo. Tendemos a dudar de la afirmación de que congelaran el agua en lugar de dudar del principio de que congelar agua la convierte en hielo.

¿Es quizá un efecto de la forma de las proposiciones? Algunos psicólogos lo creen. Supongamos que alguien nos dice:

Si estuviera lloviendo entonces Pat iría al cine.

Estaba lloviendo.

Si se plantea un conflicto, puesto que Pat no fue al cine, ¿cuál de estas dos proposiciones abandonamos si las dos están igualmente arraigadas? Supongamos que alguien nos dice:

O bien Viv se retrasó o bien no fue al cine.

Viv no se retrasó.

Si se plantea un conflicto, puesto que Viv sí fue al cine, ¿cuál de estas proposiciones abandonamos? Las pruebas psicológicas sugieren que tendemos a abandonar el condicional en el primer caso y la disyunción en el segundo caso. Lo que es extraño, sin embargo, es que los condicionales y las disyunciones son consistentes con más posibilidades que las proposiciones categóricas, y por tanto tienen una mayor probabilidad de ser verdaderos. Una razón para nuestra decisión puede ser que creemos que quienes hablan tienen menos probabilidad de equivocarse respecto a hechos categóricos sencillos. Su veracidad puede ser una cuestión de observación más que de inferencia. Por tanto, las proposiciones categóricas tienen un aura de certeza que falta en afirmaciones más complejas sobre relaciones entre proposiciones.

En suma, el arraigo de algo que se establece depende de su forma, su significado y su relación con otro conocimiento. Estos efectos no asignan ningún papel a los procesos de razonamiento mediante los que reconocemos y resolvemos las inconsistencias. Sin embargo, estos procesos tienen sus propios efectos.

* * *

Los conflictos importan. En varios experimentos, los participantes valoraron sus respectivas creencias en parejas de proposiciones. Se les enfrentó luego a hechos que eran inconsistentes con las consecuencias de esas parejas, y valoraron de nuevo sus creencias. Dependiendo de la naturaleza de la inconsistencia, a menudo abandonaron el elemento que anteriormente habían valorado como más creíble entre los que componían la pareja. El arraigo por sí solo no es una guía para la revisión de creencias.

Voy a describir un principio general de "no correspondencia" que tiene que ver con la proposición que abandonamos en caso de conflicto con un hecho. Si el hecho es inconsistente con una única creencia, entonces no importa lo arraigada que esté la creencia, deberíamos abandonarla. Podríamos jurar que hay leche en el frigorífico. Abrimos el frigorífico y no la hay. Podemos estar muy sorprendidos, pero abandonamos nuestra creencia. El caso no necesita una corroboración experimental. Pero casos más sutiles, sí. Supongamos que Pat, una amiga nuestra, nos dice:

Si el presidente tiene una villa y una piscina, entonces tiene un avión.

Y otra amiga, Viv, nos dice:

El presidente tiene una villa y una piscina.

Pero sabemos que el presidente no tiene un avión. Nuestro conocimiento está en conflicto con una consecuencia de los comentarios de Pat y Viv. ¿Cuál de las dos está equivocada? La mayoría de los participantes en un experimento rechazaron

la afirmación de Pat, y por tanto quizá estaban sesgados en contra de las proposiciones complejas. Pero consideremos este otro caso:

Pat dice: El presidente tiene una villa y una piscina, o bien tiene un avión.

Viv dice: El presidente tiene una villa y una piscina.

Pero sabemos que el presidente tiene un avión. ¿Cuál de las dos está equivocada? Ahora, la mayoría de los participantes rechazaron la afirmación de Viv. No se produjo un sesgo general en contra de las proposiciones complicadas, entonces, ¿qué pasa?

La teoría de modelos predice las respuestas, porque también en esta tarea deberíamos ser susceptibles al razonamiento ilusorio. Cuando una proposición tiene al menos un modelo mental que está en conflicto con el hecho inequívoco, la proposición parece "no corresponderse" con ese hecho. Puede no haber inconsistencia real, porque una posibilidad no representada en los modelos mentales de la proposición puede ser compatible con ese hecho. Pero, para descubrir esta posibilidad, tenemos que ampliar nuestros modelos mentales a modelos completos. Tendemos a no hacerlo, y por tanto dudamos de la proposición. En el primer ejemplo planteado anteriormente, los participantes sabían que el presidente no tiene un avión, y este hecho está en conflicto con el único modelo mental explícito de la proposición condicional de Pat. Este modelo representa que el presidente tiene una villa, una piscina y un avión.

En el segundo ejemplo, los participantes sabían que el presidente tiene un avión, y este hecho encaja con uno de los modelos explícitos de lo que tiene el presidente de acuerdo con la proposición disyuntiva de Pat: el presidente tiene una villa y una piscina, o bien tiene un avión. Por tanto, el hecho es consistente con esta proposición. Pero no está representado en el modelo de la proposición de Viv según la cual el presidente tiene una villa y una piscina. Y por tanto los participantes tienden a rechazar la proposición de Viv.

Hasta aquí, la no correspondencia podría presentarse entre la forma gramatical

de las oraciones. Por ejemplo, “el presidente no tiene un avión” no encaja con “el presidente tiene un avión”. Pero el efecto depende, no de la forma, sino de los modelos. Consideremos, por ejemplo, estas afirmaciones:

Pat dice: “Si el presidente tiene una villa y una piscina entonces o bien tiene un avión o bien un Ferrari.”

Viv dice: “El presidente tiene una villa y una piscina.”

Pero sabemos que tiene tanto un avión como un Ferrari. Deberíamos percibir la inconsistencia, y el principio de no correspondencia predice lo que sucede entonces: retiramos la proposición de Pat. Pero no podemos hacerlo sobre la base de la forma gramatical. Las dos cláusulas respecto a los hechos –tiene un avión, tiene un Ferrari– encajan con las dos cláusulas de la proposición de Pat. Tenemos que entender que el que se dé su conjunción en la realidad es inconsistente con la disyunción excluyente de la proposición de Pat. Una mera no correspondencia gramatical entre “y” y “o bien” no es bastante para establecer una inconsistencia, porque una conjunción es consistente con otras conectivas, como por ejemplo una disyunción incluyente. Por tanto, tenemos que inferir la inconsistencia. Y esta tarea, como he sostenido, se basa en modelos mentales.

* * *

La resolución de las inconsistencias no es solo cuestión de cambiar nuestras creencias. En primer lugar queremos entender qué da lugar a la inconsistencia. Eso nos ayudará a decidir qué hacer. La lista de posibilidades casi siempre es larga. Sin embargo, si podemos, eliminamos las causas probables antes de ocuparnos de las improbables. El proceso de estrechar la lista puede dar lugar a una posibilidad irresistible, pero con frecuencia dará lugar a alternativas que

compiten entre sí. Algunas veces, puede no conseguir dar absolutamente ninguna explicación –de nuevo el efecto Mary Celeste (véase el Capítulo 14)–. Un proceso importante en la resolución de la inconsistencia es por tanto el intento de pensar en una posible situación causal –un diagnóstico– que dé sentido a lo que sucede.

El proceso es abductivo. Razonamos hacia atrás a partir de una inconsistencia hasta llegar a una causa que la explique. En capítulos anteriores, describí cómo las mejores explicaciones son las cadenas causales –una causa y un efecto– y cómo utilizamos el conocimiento general para construir estas cadenas, un proceso que se simula en un programa de ordenador. El mismo mecanismo puede explicar la abducción de soluciones causales para las inconsistencias. Como ejemplo, consideremos la inconsistencia:

Si alguien tiró del gatillo, entonces la pistola disparó.

Alguien tiró del gatillo.

Pero la pistola no disparó.

Una explicación mínima sería:

No había balas en la recámara.

Pero una explicación mejor proporciona un motivo en función de una causa:

Una persona prudente descargó la pistola y no había balas en la recámara.

Por supuesto, se podría pedir también una explicación para esta causa, pero cuanto más larga es una cadena causal, más improbable se vuelve, y por tanto la mejor explicación deberían ser una causa y un efecto.

Esta hipótesis sobre las mejores explicaciones viola una asunción habitual. Los filósofos han sostenido durante mucho tiempo que los cambios en las creencias como resultado de una inconsistencia deberían ser tan mínimos como fuera posible. William James escribió: “[El nuevo hecho] conserva la vieja reserva de verdades con un mínimo de modificación, forzándolas solo lo necesario para hacerles admitir la novedad.” Teóricos más recientes han estado de acuerdo; pero si la teoría de modelos es correcta, este principio es erróneo. No nos limitamos a subir a bordo a una única proposición nueva –un efecto–, también damos alojamiento a su causa.

Un programa que implementa esta explicación utiliza los modelos para detectar una inconsistencia, y luego utiliza el principio de no correspondencia para rechazar una premisa. En el ejemplo de la pistola, el resultado es un modelo de esta situación:

Alguien tiró del gatillo y la pistola no disparó.

Este modelo provoca varias posibilidades en la base de conocimientos del programa, y el programa hace una elección arbitraria entre ellas, y extiende este modelo de manera adecuada:

La pistola estaba rota y alguien tiró del gatillo y la pistola no disparó.

De acuerdo con la teoría, la nueva conjetura sigue necesitando explicación. El programa construye esa causa, utilizando el modelo anterior para buscar en su base de conocimientos una posibilidad que dé cuenta de la ruptura de la pistola.

Entre su conocimiento está la noción de que si alguien deja caer una pistola esta puede romperse, y por tanto al modelo de la situación se le añade la cláusula si.

Alguien dejó caer la pistola y la pistola se rompió y alguien tiró del gatillo y la pistola no disparó.

El programa ha construido una cadena causal de las que, según postula la teoría, deberían resolver las inconsistencias. El programa trabaja también con modelos completos de las premisas, los cuales dan lugar a una resolución de la inconsistencia diferente.

Diversos experimentos han confirmado que los participantes prefieren explicaciones que van más allá del mínimo. En un estudio preliminar, los participantes trataron de resolver un conjunto de veinte problemas diferentes correspondientes a diversos campos. Cada problema se basaba en una inconsistencia flagrante, como la que he citado antes respecto a la pistola. La tarea de los participantes era explicar con sus propias palabras qué había sucedido. Cada problema suscitaba una variedad de explicaciones diferentes, con una media de casi cinco explicaciones por problema. Sin embargo, como predice el principio de no correspondencia, la gran mayoría de explicaciones eran equivalentes a retractaciones respecto a la proposición condicional (noventa por ciento de las pruebas) en lugar de respecto a la proposición categórica. En el dos por ciento de las pruebas, los participantes fueron incapaces de llegar a una explicación (el fenómeno Mary Celeste).

En un segundo experimento, un nuevo grupo de participantes tenía que evaluar conjuntos de explicaciones putativas para las veinte inconsistencias. Su tarea era seleccionar la explicación más probable, luego la siguiente más probable, etc., hasta que de hecho hubieran ordenado las explicaciones de acuerdo con su probabilidad. Las explicaciones se construyeron a partir de las recogidas en el estudio previo. Como predice la teoría de modelos, los participantes mostraron una abrumadora tendencia a evaluar como la explicación más probable una causa y un efecto. Por ejemplo, pusieron las explicaciones putativas para que la pistola no consiguiera disparar en este orden, empezando por la explicación más probable:

1. Una persona prudente había descargado la pistola y no había balas en la recámara. (La causa y el efecto.)

2. Una persona prudente había descargado la pistola. (La causa sola.)

3. No había balas en la recámara. (El efecto solo.)

4. La persona no tiró realmente del gatillo. (Rechazo de la proposición categórica.)

5. La pistola era pesada y no había balas en la recámara. (Conjunción no causal.)

La teoría predijo también que una causa aislada debería ser más convincente que un efecto aislado, porque el conocimiento hace más fácil inferir un efecto a partir de una causa que viceversa. La causa y el efecto forman una conjunción, y por eso se incluyó la conjunción no causal como control. Una debilidad del experimento era que todos los participantes recibían los veinte guiones en el mismo orden. Sin embargo, en un experimento posterior en el que se presentaron ordenados de cuatro formas diferentes se produjeron los mismos patrones de resultados. Ocurrió también en un experimento en el que las explicaciones putativas excluían la premisa categórica, p. ej.:

La persona estaba semiparalizada y no fue capaz de mover los dedos con suficiente fuerza.

No siempre acomodamos una nueva creencia haciendo un cambio tan pequeño como sea posible en nuestras viejas creencias. La aceptación de una conjunción exige un cambio mayor que la aceptación de una sola de sus proposiciones. La preferencia por la conjunción sorprende en otro sentido. El que se la valore como más probable que cualquiera de sus componentes es contrario a las normas de probabilidad: una conjunción no puede ser más probable que uno de sus componentes. La mejor explicación no es siempre la proposición mínima y más

probable.

* * *

¿Cómo influye el razonamiento en nuestras creencias? La inferencias inconscientes nos llevan a tener un lazo emocional con las creencias sobre cosas que no hemos visto. Estas creencias son vitales para nosotros, y nuestro lazo con ellas con frecuencia las hace inmunes a un argumento razonado. Nuestros estereotipos pueden ser benignos –la mayoría de los perros tiene cuatro patas–. Pero también pueden ser malignos –las mujeres no son peores conductoras que los hombres–. Cuando forman parte de nuestra visión del mundo, pueden ser difíciles de cambiar. Nuestra creencia en ellas puede inhibir nuestra búsqueda de contraejemplos para inferencias que las apoyan. Sin embargo, sí que modificamos nuestras creencias prosaicas, en caso de un choque con los hechos. Abandonamos las creencias que están en conflicto con los hechos, o las que parecen estarlo –un parecido que se puede explicar por una no correspondencia con los modelos mentales–. Intentamos razonar una manera de resolver la inconsistencia. Nuestra capacidad para hacer estas abducciones es muy superior a cualquier programa de ordenador existente, pero no es, y no puede ser, perfecta. De vez en cuando, somos incapaces de crear una explicación que resuelva una inconsistencia. Y a menudo modificamos nuestras creencias para aceptar en ellas cambios que no son todo lo pequeños que podrían ser. A pequeña escala, nosotros también seguimos el camino de Damasco.

CAPÍTULO 24

Cómo resolvemos los problemas

Supongamos que usted ha pintado un cuadro del famoso Baptisterio de Florencia, y quiere comprobar su exactitud. ¿Cómo lo haría? Cuando planteé esta pregunta a un grupo de estudiantes, sugirieron que deberían tomar medidas del edificio y comparar sus proporciones con las de su cuadro. Esto serviría para un dibujo de la fachada de un arquitecto, pero su cuadro refleja cómo se ve el edificio desde un determinado punto, y líneas que en el edificio son paralelas en un cuadro deben converger. Por tanto, ¿cómo podría usted comprobar que las ha hecho bien? Los estudiantes sugirieron compararlo con una fotografía tomada desde el mismo punto. Bastante aceptable; pero supongamos que el cuadro lo pintó usted en el siglo quince mucho antes de la invención de la fotografía. Este problema se le planteó al gran artista y arquitecto del Renacimiento Filippo Brunelleschi, y lo resolvió con un insight extraordinario. Hizo un agujerito en el centro de su cuadro, puso el cuadro mirando hacia el edificio y miró a través del agujero. Vio cómo se veía el edificio desde ese punto. Y luego colocó un espejo frente a él entre el cuadro y el edificio de manera que le permitiera ver reflejado el cuadro a través del agujerito. Colocando el espejo dentro y fuera de su línea de visión, podía hacer una comparación casi inmediata entre el cuadro y el propio edificio. De esta forma, Brunelleschi descubrió los principios de la perspectiva y el uso de puntos “de fuga”.

No sabemos cómo ideó Brunelleschi este gran truco con un espejo, pero imagino que luchó con el problema y luego tuvo un soplo de inspiración a partir del cual la solución estalló en su mente como si viniera de ninguna parte. Todos hemos tenido ese tipo de experiencias. Decimos: “Eureka, lo tengo”, como se supone que hizo Arquímedes cuando imaginó cómo medir el volumen de la corona del rey de Siracusa. Nosotros, sin embargo, no saltamos del baño y corremos desnudos por las calles. Ese insight parece diferente del proceso normal de pensamiento, en el que avanzamos paso a paso hacia la solución de modo similar a cuando utilizamos un mapa para averiguar cómo ir de un sitio a otro. Ciertamente, en psicología hay una larga tradición respecto a que el insight es

excepcional. La experiencia es especial, ¿pero depende de un tipo especial de pensamiento? Los escépticos lo dudan. Si buscamos el número que abre una cerradura de combinación, su descubrimiento necesariamente es repentino, pero no depende de un proceso que se diferencie de los pasos previos que no lo consiguieron. Simplemente probamos una nueva combinación, y funciona. Sin embargo, algunas pruebas sugieren que el insight se diferencia de la monótona aplicación de los métodos habituales. Cuando actuamos de esta forma rutinaria, nos damos cuenta de cuándo nos acercamos a la solución. Pero un insight no se presagia mediante ninguna sensación de aumento de "temperatura". En un determinado momento no tenemos ni idea de la solución; en el momento siguiente la hemos conseguido. Incluso las pruebas de estudios con imágenes cerebrales sugieren que en el momento de producirse un insight se activan partes especiales del cerebro. Sin embargo, nadie ha conseguido precisar qué se considera un insight y qué procesos mentales subyacen en él. Por tanto, dejo a un lado estos misterios por un tiempo. En vez de eso, consideraré el razonamiento, y su utilización en la resolución de problemas. Después volveré al insight.

Quienes estudian la creatividad con frecuencia sostienen que somos demasiado lógicos como para ser todo lo creativos que podríamos ser. Necesitamos dejar de pensar de una forma mojigata. Necesitamos pensar "lateralmente" y "de forma innovadora". Quienes estudian el razonamiento, por el contrario, sostienen que somos demasiado creativos como para ser todo lo lógicos que podríamos ser. Necesitamos dejar de pensar de forma indisciplinada y empezar a pensar más "linealmente". La verdad del asunto es que la imaginación nos ayuda a razonar, y el razonamiento nos ayuda a imaginar. Los teóricos pueden trazar una línea entre la deducción y los otros tipos de razonamiento: la deducción da lugar a conclusiones que deben ser verdaderas dadas las premisas. Pero el razonamiento abductivo es un ejercicio de imaginación. Y por tanto el presente capítulo esboza una teoría respecto a cómo el razonamiento, tanto deductivo como inductivo, nos ayuda a resolver los problemas. Es un avance del análisis de la pericia que aparece en la parte final del libro.

* * *

¿Qué significa decir que un proceso es creativo? Para conocer un melón hay que

catarlo. Examinemos los resultados. Deberían ser novedosos –como mínimo no deberíamos haber recordado la misma idea que ya hemos tenido antes y, con suerte, podrían incluso ser novedosos para el mundo en general–. Esta es una opción, porque si hemos pensado en la idea nosotros solos, seríamos creativos aunque alguien más se nos hubiera anticipado. Quizá esta noción subyace en la extraña historia de Borges Pierre Menard, autor del Quijote, en la que un autor simbolista francés está tratando de reescribir El Quijote de Cervantes palabra a palabra, no plagiándolo, sino como resultado de su propio proceso creativo. Como sostuve en capítulos anteriores, el pensamiento no es determinista, y por tanto el proceso creativo tampoco lo es. La abducción puede crear ideas con un contenido proposicional, es decir, podemos preguntar si los resultados son verdaderos o falsos. Otros tipos de creación, como la composición musical, dan lugar a productos que no tienen un contenido proposicional.

Cuando hacemos abducciones, trabajamos dentro de un paradigma reconocido por la ciencia o dentro de un género reconocido por el arte. Creamos una hipótesis científica o una obra de arte. Como consecuencia, el proceso está restringido por nuestro conocimiento respecto a ese campo, aunque podemos no ser conscientes de gran parte de este conocimiento. Nuestra mente contiene constricciones inconscientes, que discutí en capítulos anteriores, y que gobiernan nuestros esfuerzos para crear. Somos más o menos tan conscientes de ellas como de las reglas de gramática de nuestra lengua materna. Es decir, somos conscientes de algunas reglas –por ejemplo, el sujeto de una oración debe concordar en número con el verbo–. Pero somos ajenos a la mayoría de ellas, razón por la que existe la disciplina de la lingüística. De la misma forma, somos conscientes de algunas constricciones que gobiernan un paradigma o un género. La velocidad de la luz es constante. Los genes codifican las proteínas. Un soneto tiene catorce versos. Pero ignoramos muchas constricciones que nos guían en el proceso de abducción. Si fuera de otra forma, sería fácil programar un ordenador para imitar nuestros pensamientos, o para probar que el proceso no es computable. Aquello de lo que somos conscientes es el conjunto de elementos que reunimos para crear. La ciencia está hecha de conceptos, como por ejemplo genes y proteínas; la música está hecha de notas, y los poemas –como Malearé recordó al pintor Degas– están hechos, no de ideas, sino de palabras.

Con esto he descrito los cinco componentes de una definición de partida para la creatividad, la llamada definición “ONCE”: la N ovedad para nosotros, la novedad O pcional para la sociedad, un proceso N o determinista, las C onstricciones del paradigma o género y la utilización de E lementos existentes.

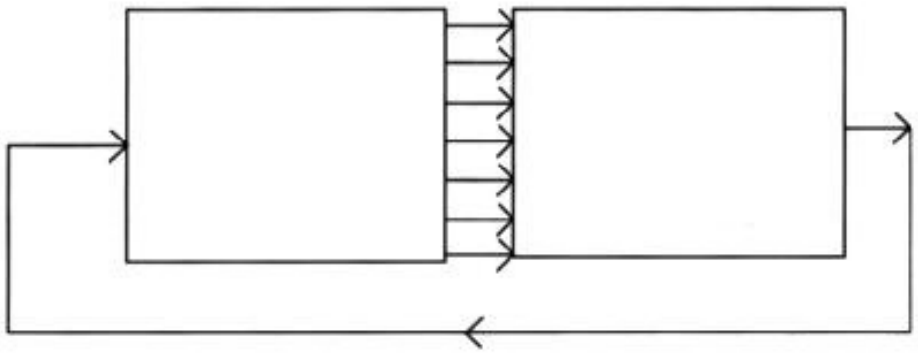
La definición delimita lo que se considera como creativo, pero peca de generosa. Hasta ahora, sin embargo, nadie ha encontrado un caso de creatividad genuina que quede fuera de ella. El candidato más sólido es el descubrimiento de una prueba matemática, pero quienes recurren a él confunden el determinismo evidente de una prueba con el muy diferente proceso de su descubrimiento –incluso los programas de ordenador que crean pruebas simulan procesos que son no deterministas–. Dada la definición ONCE y la asunción de que la abducción es un proceso computable, se sigue que solo hay tres tipos de procesos computacionales que podrían crear nuevas ideas.

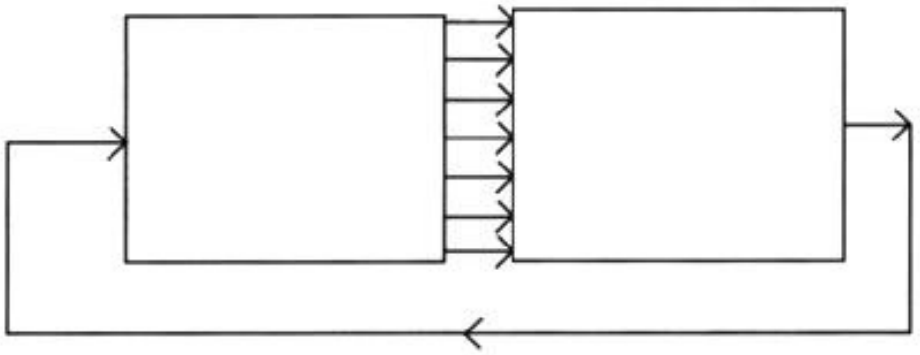
El primer proceso imita la evolución de las especies. Como los genes, los conceptos existentes se mezclan al azar para generar nuevas ideas. Como la selección natural, las constricciones basadas en el conocimiento y la experiencia seleccionan aquello que es viable –si existe algo– y a lo que se puede aplicar el proceso completo una y otra vez hasta que tal vez surja algo útil. Actualmente los programas de ordenador simulan esta evolución de las ideas, y pueden aproximar soluciones óptimas para ciertos problemas, como por ejemplo minimizar la longitud de las conexiones en un circuito integrado. La mayoría de los psicólogos que han pensado sobre la creatividad defienden para ella esta descripción “neodarwinista”. Pero el proceso no podría ser más ineficiente. La evolución funciona de esta forma solo porque las experiencias de un organismo individual no se pueden codificar en los genes que pasa a sus descendientes. Nuestra creatividad rara vez ese tipo de ensayo y error, en el que el conocimiento actúa como un mero tamiz.

El segundo proceso es más eficiente. Utiliza algunas constricciones en la generación de ideas. Otras se utilizan para evaluar los resultados y hacer modificaciones en ellos, y los resultados pueden retroalimentar el estadio generativo para un trabajo posterior. Puesto que las constricciones gobiernan la generación, el procedimiento ya no es simplemente ensayo y error, sino que se ajusta a la tarea habitual de generar las ideas iniciales, revisarlas, revisarlas de nuevo, etc., en múltiples borradores. Este tipo de creación es una cuestión de “múltiples estadios”. Su estadio generativo, como todos los procesos mentales, es inconsciente. El estadio evaluativo puede aplicar constricciones conscientes.

El tercer tipo de proceso lleva el procedimiento de múltiples estadios al extremo. Todas nuestras constricciones gobiernan la generación de ideas. Por tanto, si las constricciones son viables, el resultado no necesita revisión. Por definición, no existen más constricciones para evaluarlo. En diversos puntos del procedimiento,

puede haber más de una opción, porque las constricciones rara vez dan lugar a que sea posible un único paso. Pero, dado que hemos utilizado todas las constricciones que tenemos, la elección entre las opciones debe ser arbitraria. Es en la elección donde entra en el proceso el no determinismo de la creatividad. Es "neolamarckiana" por analogía con la teoría de la evolución de Lamarck, porque todas las constricciones que hemos adquirido a partir de la experiencia acuden a gobernar el estadio generativo. Con un conjunto viable de constricciones un proceso neolamarckiano es muy eficiente. Nos permite crear en tiempo real, y por tanto subyace en cualquier tipo de arte en el que se improvise. Por ejemplo, los músicos de jazz, que improvisan para vivir, no pueden volver atrás para revisar una frase que acaban de tocar. Por tanto, necesitan desarrollar un conocimiento tácito de las constricciones del género, que pueden utilizar para crear música. Cuando dominamos el método para resolver ciertos tipos de problemas sin dar nunca un paso en falso, también hemos abducido un proceso neolamarckiano. La Figura 24.1 resume los tres tipos de proceso creativo.





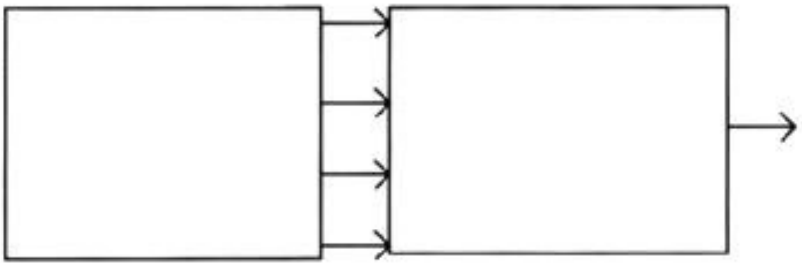


Fig. 24.1. Tres tipos de proceso creativo.

* * *

La mayoría de los problemas de la vida real exigen un proceso de estadios múltiples. Utilizamos algunas constricciones para generar una solución putativa, y otras constricciones, como el objetivo del problema, para criticar y corregir los resultados. Si nuestros esfuerzos no tienen éxito, entonces seguimos trabajando algo más sobre ellos, etc., hasta que, o bien tenemos éxito o bien abandonamos. El éxito depende de que tengamos el conocimiento requerido y seamos capaces de recordarlo. Los problemas que exigen un insight son a menudo de este tipo. Por ejemplo, en el problema de la vela (véase el Capítulo 6), tenemos una vela y una caja de chinchetas y nuestra tarea es fijar la vela en la pared sin utilizar nada más. Clavamos una chincheta en la vela, pero no sobresale lo suficiente por el otro lado como para clavar la vela en la pared. La mayoría no sabemos qué hacer al llegar a este punto. La solución depende de un insight: podemos clavar en la pared la caja que contenía las chinchetas, y utilizarla para apoyar la vela. La solución parece trivial una vez que la conocemos, pero no es fácil de descubrir. Tenemos el conocimiento requerido, pero necesitamos también el insight de que un objeto que sirve para una función puede servir también para otra. Tendemos a tener una idea fija sobre lo que un objeto puede hacer, y por tanto el problema es difícil.

Muchos problemas nos vencen porque nos falta el conocimiento para resolverlos. Cuando la televisión falla, el ordenador no carga el sistema operativo o el teléfono se corta, podemos necesitar expertos que los arreglen. Pero podemos desarrollar nuestra propia pericia si tenemos los rudimentos necesarios para encontrar una solución. Una vez que hemos resuelto el problema de la vela, retenemos nuestro conocimiento sobre su solución, y por tanto cualquier otro ejemplo del problema es trivial. Pero otros tipos de problemas no son únicos, sino que aparecen en muchas versiones cada una de las cuales exige pensar. Planear un viaje, escribir un programa de ordenador y rellenar un formulario de impuestos son todos ellos ejemplos de problemas que algunos

tenemos que resolver una y otra vez. Cuando tratamos de resolver uno de estos problemas por primera vez, no sabemos qué hacer. Damos palos de ciego intentando un paso táctico y luego otro con pocas constricciones que gobiernen nuestras opciones. No tenemos ninguna estrategia para resolver el problema, donde una estrategia es un conjunto de constricciones sistemáticas que guían nuestra generación de pasos tácticos elementales para conducirnos a una solución (véase el Capítulo 19).

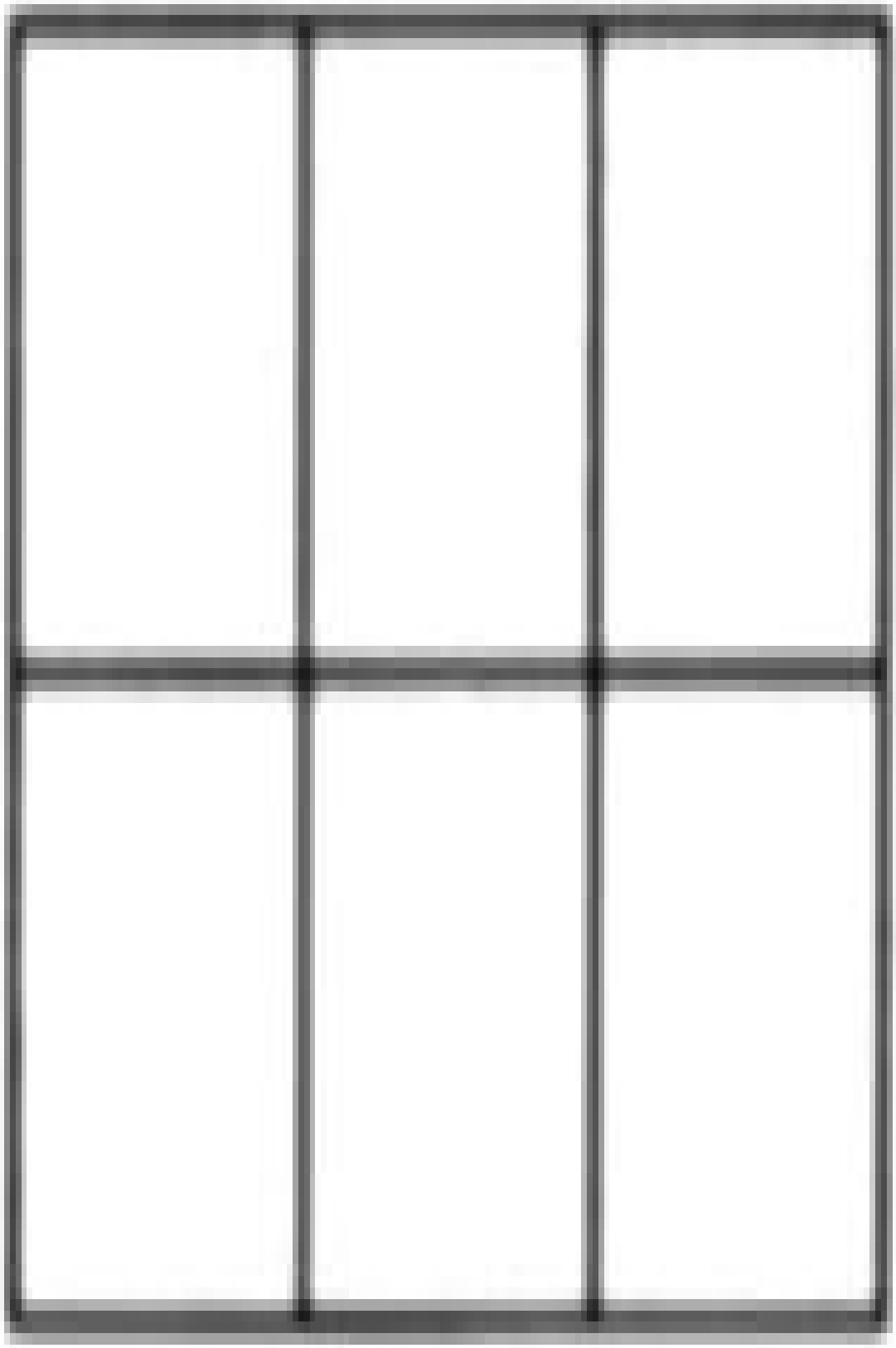
Aprendemos de nuestros errores. La idea es obvia para todos excepto para aquellos psicólogos que sostienen que se nos deberían enseñar de una forma que nos impidiera cometer errores. Cualquiera que haya aprendido a escribir programas de ordenador sabe que esta meta es tan imposible como poco aconsejable. También aprendemos de los pasos tácticos correctos. Esta idea es obvia para todos excepto para aquellos psicólogos que abogan por teorías en las que el aprendizaje se produce solo cuando nos equivocamos. Generamos un paso táctico y evaluamos sus consecuencias. Deducimos que es inútil porque no nos hace avanzar hacia una solución. O deducimos que es útil. En cualquiera de los dos casos, podemos aprender las consecuencias de ese paso táctico. Por tanto, incluso aunque el paso puede haber sido en vano y necesitemos deshacerlo más tarde, podemos haber aprendido algo que será útil en el futuro.

Una asunción clave de la teoría que yo propongo es que nuestro sistema para resolver problemas puede llevar el conocimiento desde el estadio evaluativo del proceso al estadio generativo. Por tanto, el conocimiento que deducimos respecto a un paso táctico puede trasladarse de forma que empiece a restringir nuestra generación de pasos –estamos progresando hacia un proceso neolamarckiano–. El traslado depende de la experiencia, pero nos capacita para evitar errores que más tarde tendríamos que deshacer. Vamos a ver algunos ejemplos que apoyan esta teoría.

* * *

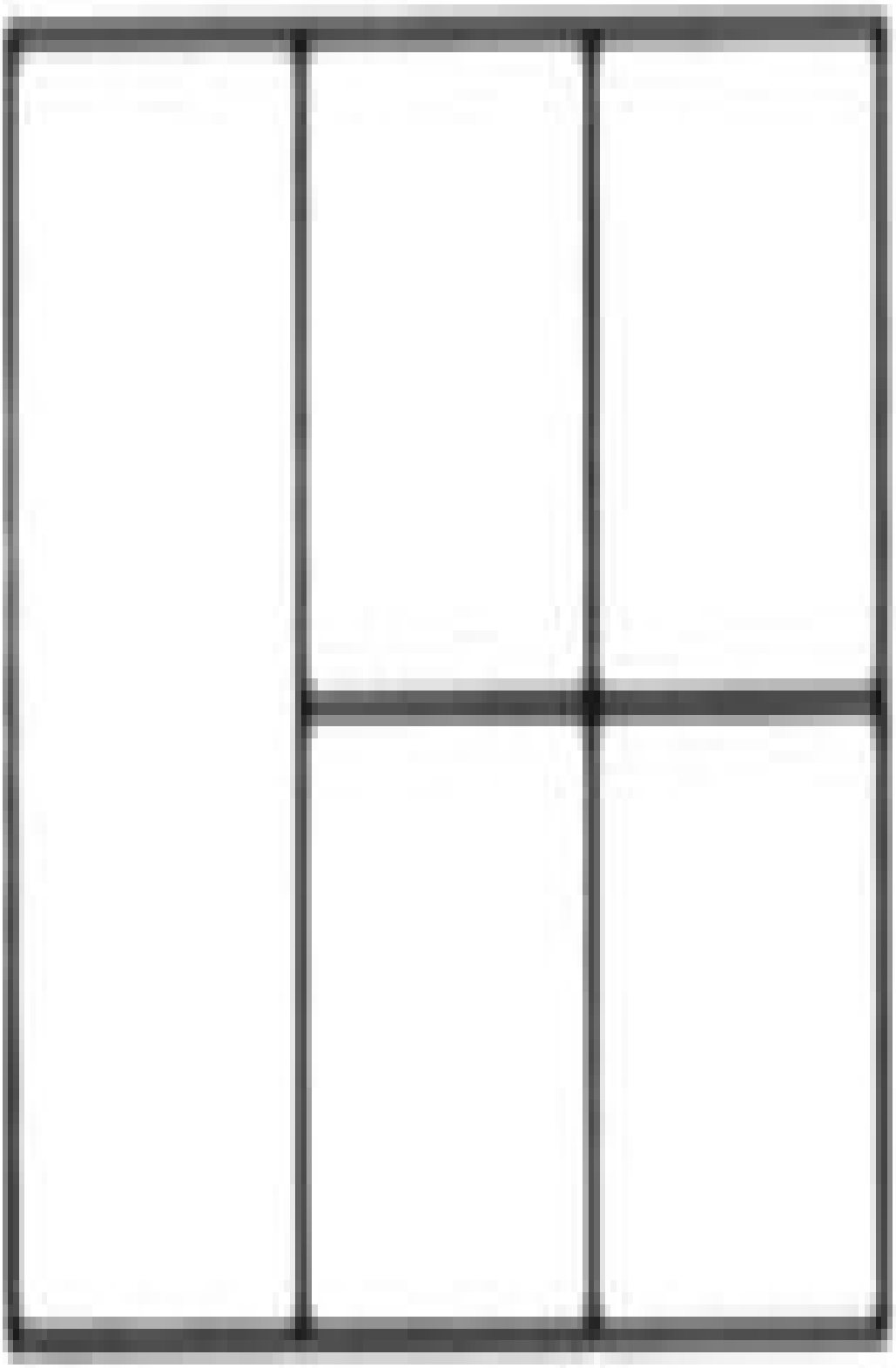
Los problemas de “forma” nos exigen modificar una forma añadiendo, quitando o reorganizando alguna de las piezas que la componen. Ilustran muchos de los rasgos respecto a cómo resolvemos problemas que en general no exigen un

insight, y este es un ejemplo del tipo de problemas en el que me voy a centrar. Dados los seis cuadrados que aparecen a continuación, quitar cinco piezas para dejar exactamente tres cuadrados sin extremos sueltos, es decir, sin piezas que no estén unidas a ninguna por algún extremo:



Le invito a tratar de resolver este problema por sí mismo ya sea mentalmente o utilizando un modelo externo hecho con cerillas o palillos de dientes.

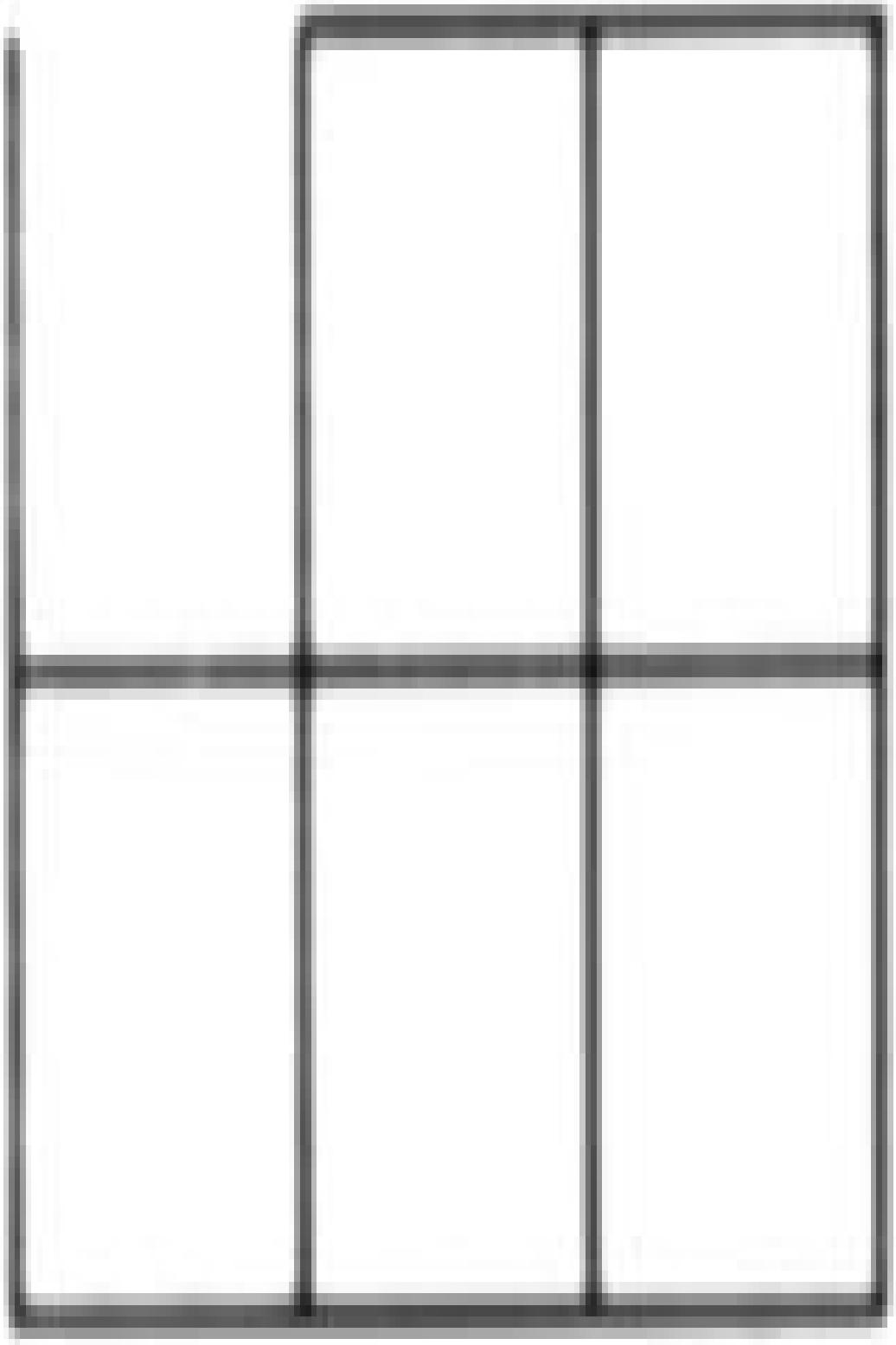
Cuando no hemos resuelto nunca antes uno de estos problemas solemos intentar uno o dos pasos experimentales utilizando una estrategia neodarwiniana. Supongamos que quitamos una pieza de la fila central:



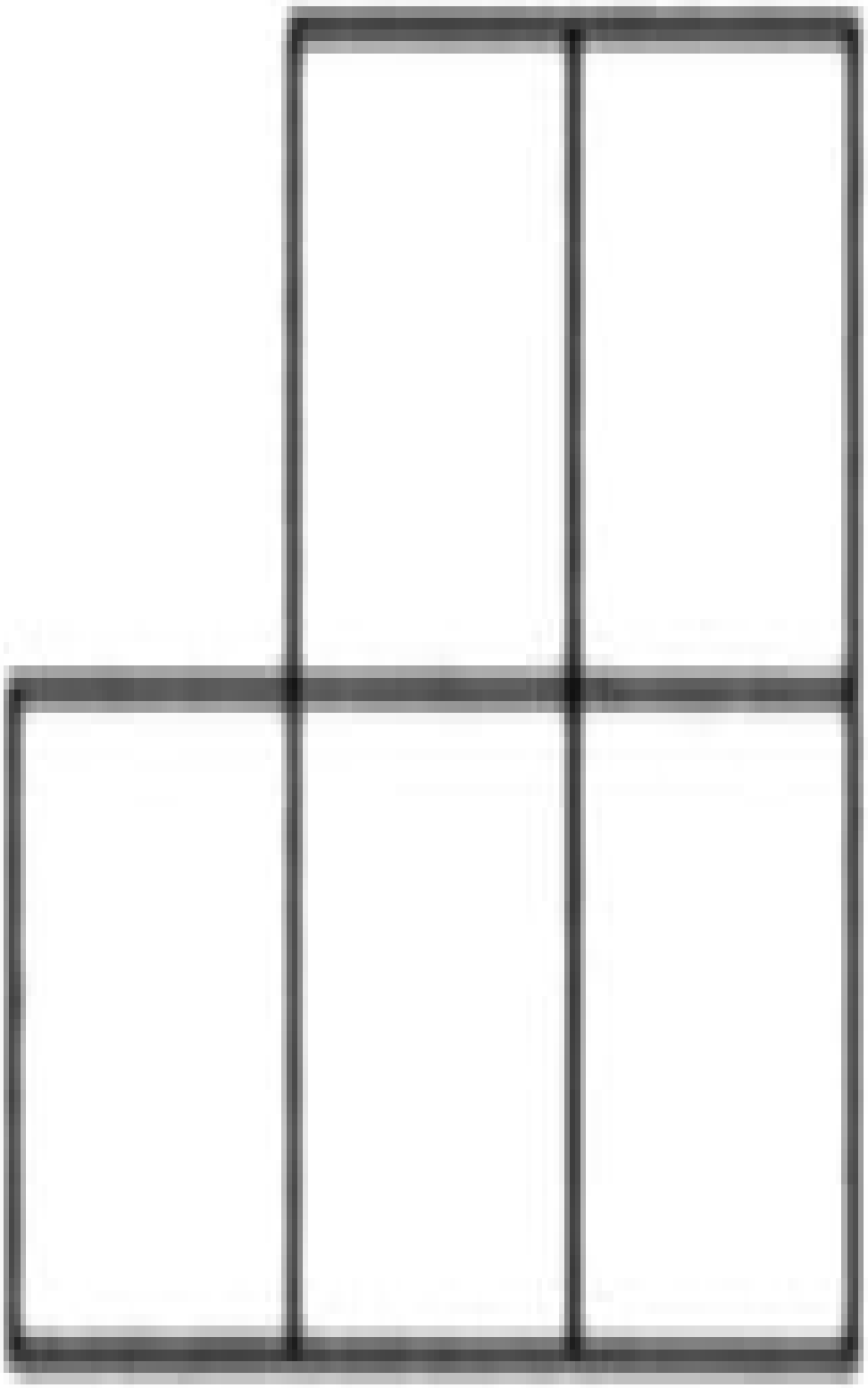
Quitamos una pieza y eliminamos dos cuadrados. Y, como podemos deducir, el paso transforma el problema en otro nuevo: ahora tenemos que quitar cuatro piezas de la distribución resultante para dejar exactamente tres cuadrados sin extremos sueltos. Por tanto, ¿nuestro paso es bueno o no? No tenemos forma de responder a esta pregunta sin llevar a cabo, quizá mentalmente, algunos pasos más y evaluar las consecuencias. Por consiguiente, no tenemos forma inmediata de evaluar si un paso es útil, y por tanto tenemos que continuar intentando pasos.

Cuando examinamos la distribución inicial de los cuadrados, vemos inmediatamente una distribución simétrica de cuadrados que forman un rectángulo. Ciertas partes de la figura son más notorias que otras, y en particular aquellas partes cuyo contorno externo cambia de dirección. Por tanto, lo más notorio son las cuatro esquinas. Tanto la simetría como la notoriedad deberían afectar a los pasos que damos, porque nos centramos en las piezas notorias, y la simetría nos anima a buscar soluciones simétricas. El programa de ordenador que escribí para aplicar un proceso neodarwiniano a estos problemas realiza el trabajo en serie de una forma metódica y lenta para identificar los cuadrados de la distribución. Tiene una plantilla de un cuadrado, e identifica los cuadrados comparando esta plantilla con cada célula de la distribución. Sitúa la plantilla sobre una de las celdas, y comprueba, una a una, si cada pieza de la plantilla tiene una pieza que le corresponda en la distribución. Si es así, ha identificado un cuadrado, y se traslada a la siguiente celda. Identifica otras formas, como esquinas de cuadrados, utilizando de forma similar plantillas que se ajusten a ellas, aunque también puede girar la plantilla. Por el contrario, nuestra vista lleva a cabo un gran número de operaciones de este tipo en paralelo, y por eso es por lo que de un vistazo vemos una distribución de cuadrados. De lo que no nos damos cuenta en un primer momento es de que algunas piezas forman parte de dos cuadrados.

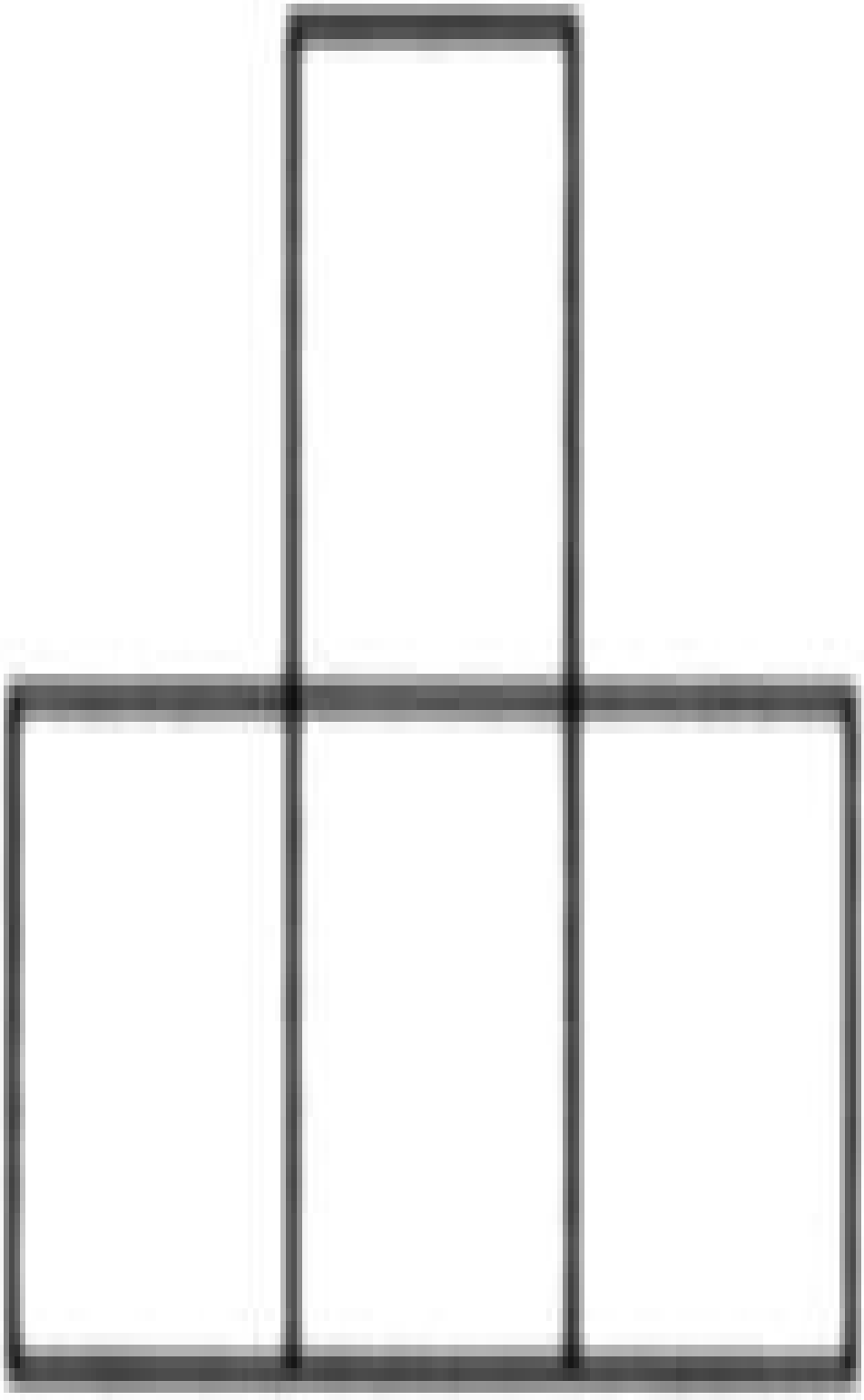
A medida que ensayamos pasos, aprendemos sus consecuencias tácticas. Nuestro paso previo, por ejemplo, eliminó una pieza intermedia que formaba parte de dos cuadrados de manera que quitamos una pieza pero eliminamos dos cuadrados. Si decidimos que el paso era un callejón sin salida, recolocamos la pieza e intentamos un paso diferente:



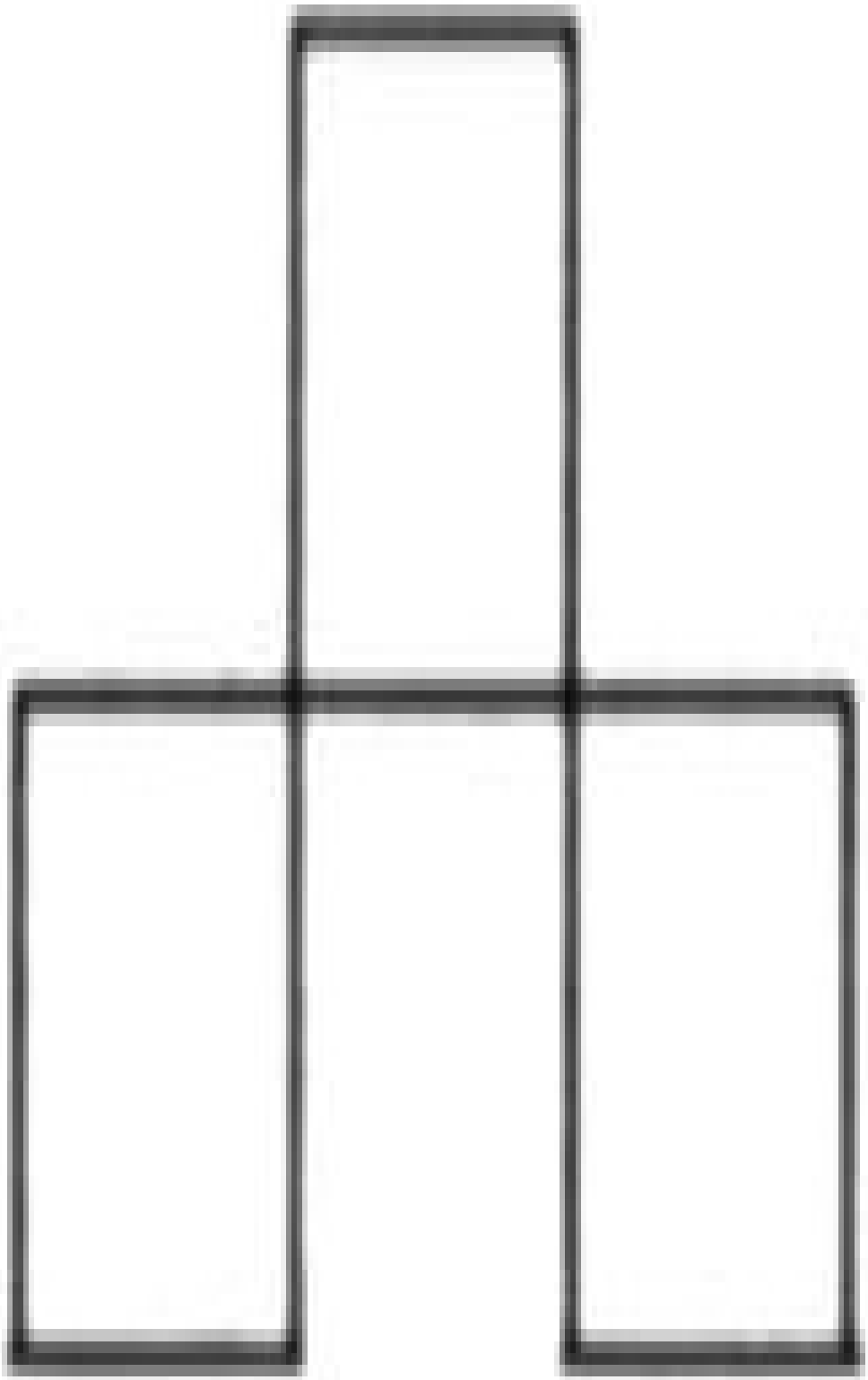
Ahora tenemos que quitar el extremo suelto que queda para dar lugar a:



Aprendemos que otra forma de eliminar un cuadrado es quitar las dos piezas que forman una esquina. Podemos deducir que ahora tenemos que quitar tres piezas y eliminar dos cuadrados. Supongamos que quitamos otra esquina:



Deducimos que ahora tenemos que quitar una pieza para eliminar un cuadrado. Podemos tener éxito:



Una vez más, aprendemos que podemos eliminar un cuadrado quitando una pieza exterior, es decir, una pieza de un cuadrado que en ambos extremos tiene piezas que pertenecen a dos cuadrados, pero que ella misma no pertenece a dos cuadrados. Con experiencia suficiente, aprendemos los efectos de los otros pasos tácticos: podemos quitar un cuadrado aislado, una U formada por tres piezas, o una unión que no forma parte de un cuadrado pero está conectada por ambos extremos con otras piezas.

Las grabaciones en vídeo de los participantes mientras resolvían estos problemas mostraron que realmente adquirían conocimiento táctico –daban menos pasos en falso a medida que aumentaba su experiencia–. Utilizaban estrategias en las que se enfrentaban a cada problema en dos fases. La primera fase era exploratoria: intentaban diversos pasos tácticos, que posteriormente tenían que deshacer. Pero estaban aprendiendo los efectos de los diversos pasos tácticos, y deduciendo las consecuencias para avanzar hacia una solución. La duración de esta fase dependía de la experiencia de los participantes, y por tanto caía proporcionalmente a lo largo de los problemas a medida que adquirían conocimiento. La segunda fase de la estrategia era la aplicación abductiva del conocimiento táctico. Un factor clave era la relación entre el número de piezas a retirar y el número de cuadrados a eliminar. Los participantes deducían esta relación y la utilizaban para restringir la selección de los pasos tácticos apropiados. Sin embargo, su utilización óptima depende del conocimiento de todos los diferentes pasos tácticos. Los participantes captaban su relevancia, pero pocas veces de forma completa.

* * *

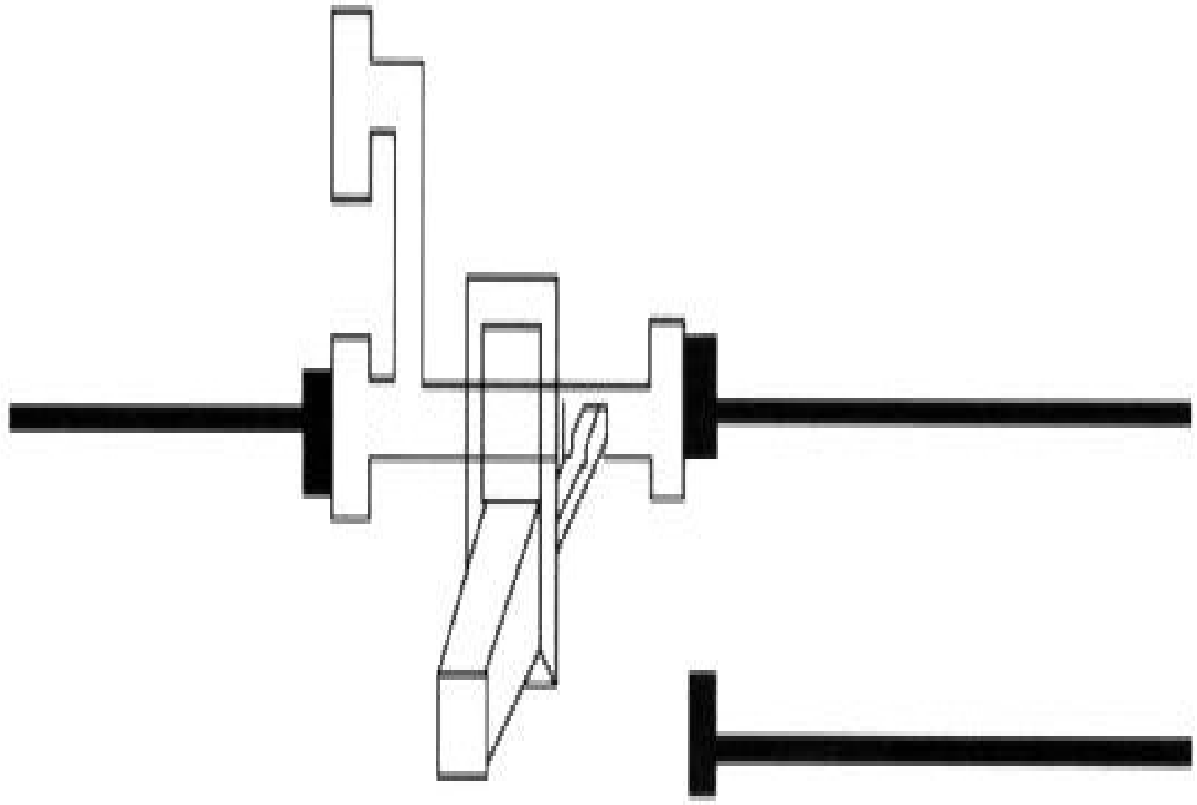
Supongamos que usted trabaja para una compañía que fabrica hornos tostadores, y una compañía rival ha puesto en el mercado un ingenioso nuevo modelo. Su jefe le presenta uno de ellos, y dice, “Averigüe cómo funciona, para que podamos hacer algo similar sin infringir su patente”. Su tarea es desmontar la cosa, y utilizar su conocimiento de cómo funcionan los componentes para inferir, mediante una serie de abducciones, cómo funciona el aparato como un todo.

Tiene usted que construir un modelo que simule sus operaciones. También puede que además tenga que descubrir cómo funcionan los componentes. Se montaron para satisfacer la pormenorización de lo que debería ocurrir cuando la persona que utiliza el aparato selecciona diversas posiciones de los controles. Puede usted informarse sobre estas posiciones en el manual de instrucciones y tiene que trabajar hacia atrás, a partir de ellas, para inferir cómo implementa el aparato esas diversas funciones. Tiene que abducir posibles formas de unir los componentes, utilizando un procedimiento generativo restringido por el conocimiento del que dispone. Luego evalúa los resultados deduciendo las consecuencias del sistema que usted ha montado. Compara el resultado con el requerido, y de esta forma deduce si el aparato está haciendo lo que se supone que hace. En resumen, su tarea es hacer "ingeniería inversa" sobre el horno tostador. Y mi tarea es hacer ingeniería inversa sobre cómo lo hace usted.

* * *

¿Qué tipo de dominio constituye un campo de pruebas adecuado para la ingeniería inversa? Los hornos tostadores y sus mandos no son ideales, porque dependen de un conocimiento demasiado especializado. Un dominio que resulta ideal –no demasiado simple y no demasiado complicado– es uno que implementa la lógica booleana, que presenté anteriormente (véase el Capítulo 8). Se refiere a la lógica del no, y y o que formalizó por primera vez el lógico del siglo diecinueve George Boole. La unidad de procesamiento de un ordenador está formada por componentes booleanos, porque también bastan para la aritmética. Un ejemplo más simple de un problema booleano es un circuito eléctrico que controla si una luz se enciende o no. Supongamos, por ejemplo, que hay dos interruptores, y la luz se enciende cuando un interruptor o el otro están pulsados, o ambos lo están. Su tarea es hacer ingeniería inversa con un circuito para este problema con "o", es decir, inferir cómo conectar los dos interruptores, la pila y la bombilla, de forma que el circuito funcione de la manera requerida. Los dos interruptores que tiene a su disposición son del tipo de los que aparecen en la Figura 24.2. Tienen un terminal en un lado y dos terminales en el otro lado, y el interruptor conecta el terminal único de un lado con el terminal superior o inferior del otro lado. Si usted conecta uno de los dos terminales del lateral del interruptor, entonces tiene un interruptor sencillo

“abierto o cerrado”; si usted conecta ambos terminales, entonces el interruptor forma parte de dos circuitos alternativos, y cuando cierra un circuito abre el otro. Solo por diversión, quizá le apetezca ver si puede resolver el problema con “o”, suponiendo que usted no sea un ingeniero eléctrico que ya conoce la respuesta. Una vez más, puede tratar de resolver el problema utilizando un modelo externo, como un diagrama o un circuito de juguete con cables y una bombilla. Puede conectar los dos interruptores de cualquier forma que desee con tal de que la luz se encienda cuando uno u otro de los interruptores está pulsado o lo están los dos.



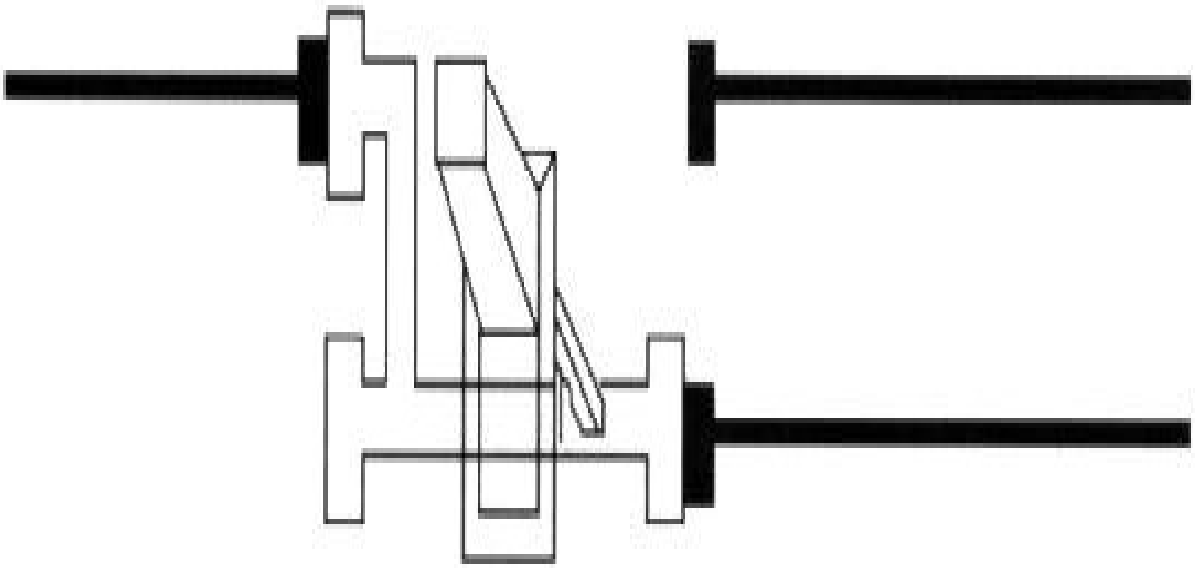


Figura 24.2. Las dos posiciones de un interruptor en un circuito booleano.

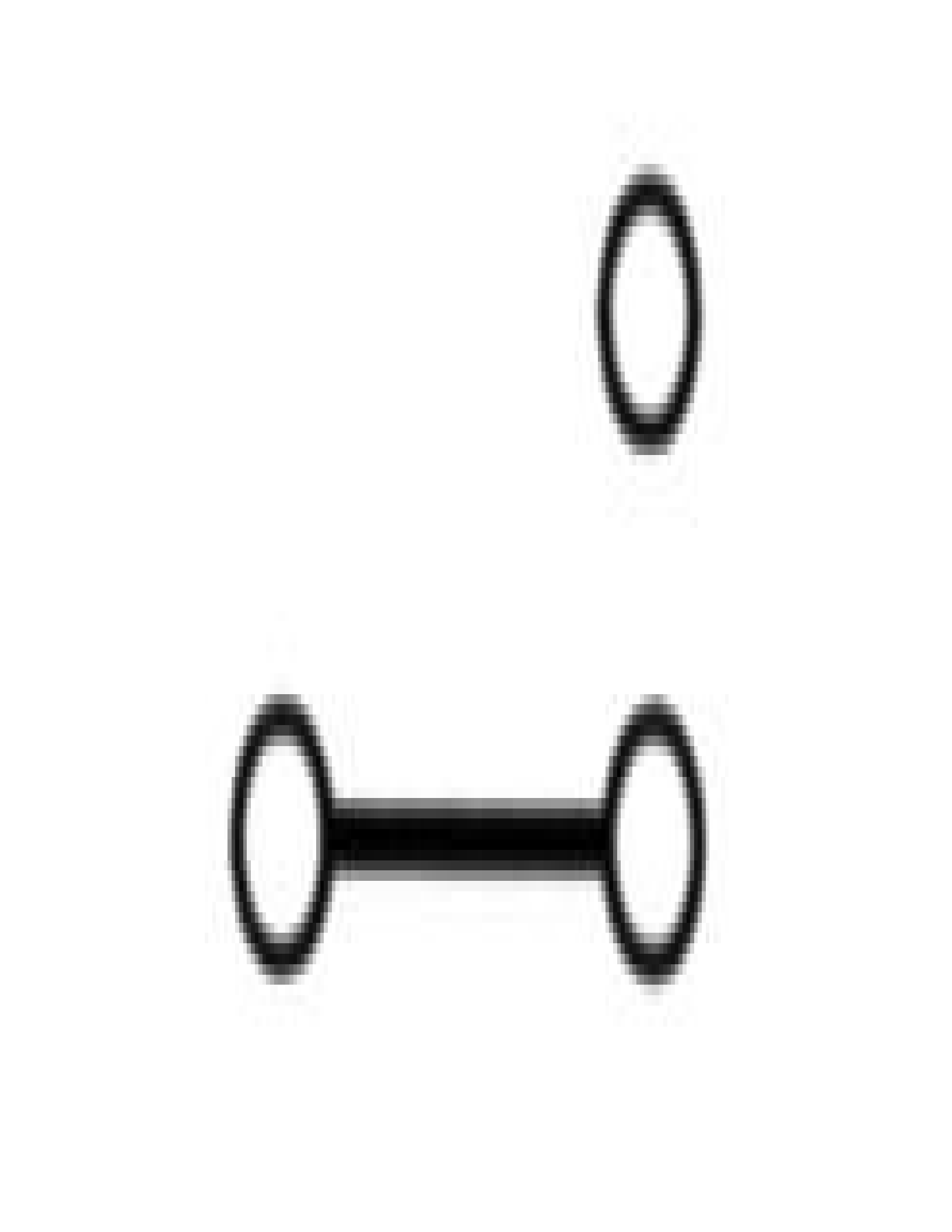
¿Cómo lo abordó? Quizá le apetezca comparar su actuación con la teoría que voy a esbozar sobre cómo resolvemos estos problemas. La teoría da por hecho que entendemos el problema –en este caso las posibles posiciones de los interruptores y los efectos resultantes sobre la bombilla–. Solo para estar seguro de que usted también entiende el problema, explicaré con detalle las cuatro posibilidades:

Interruptor 1	Interruptor 2	Bombilla
Pulsado	Pulsado	Encendida
Pulsado	Sin pulsar	Encendida
Sin pulsar	Pulsado	Encendida
Sin pulsar	Sin pulsar	Apagada

Como ve usted, he adoptado la convención europea de que la posición “encendido” de un interruptor corresponde a que esté pulsado. La teoría también asume que tenemos algún conocimiento rudimentario sobre los circuitos eléctricos y que entendemos cómo funcionan los componentes –en este caso qué hacen las dos posiciones del interruptor–. También necesitamos alguna idea de los pasos tácticos que podemos utilizar para conectar un circuito, p. ej., conectar un cable a una pila o a un interruptor.

La teoría postula que utilizamos algunas constricciones para abducir posibles

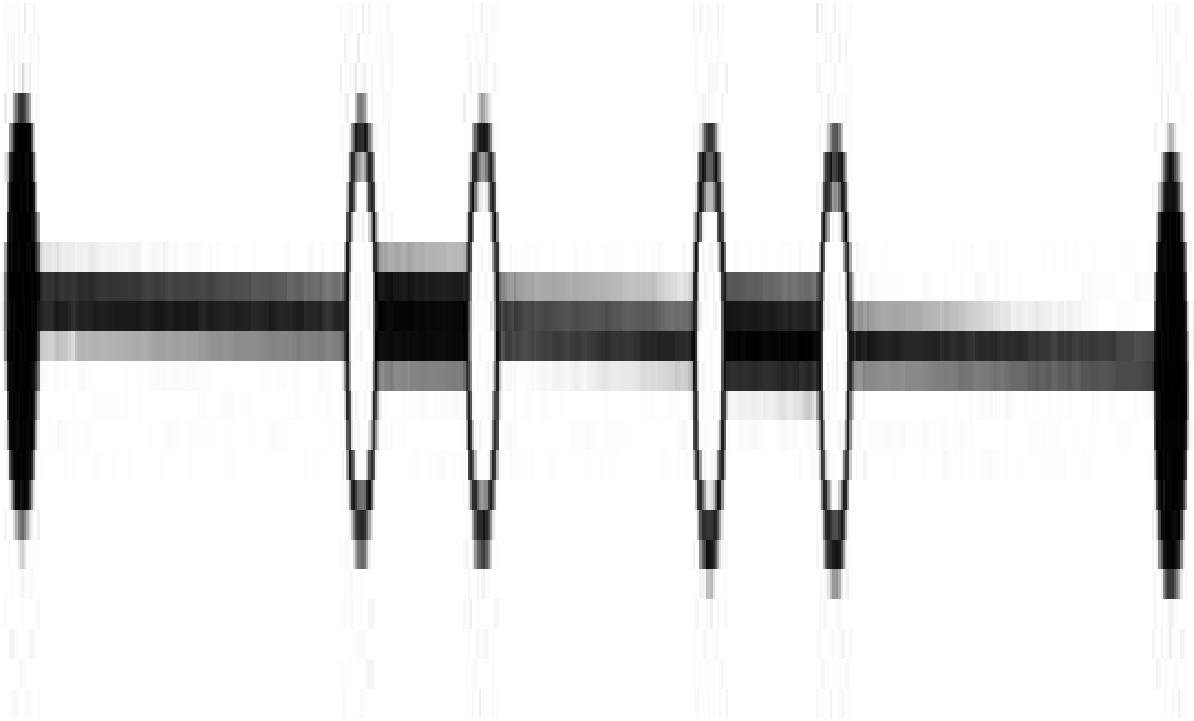
pasos hacia una solución y otras constricciones para deducir las consecuencias de estos pasos. Es decir, utilizamos un proceso de estadios múltiples. El primer estadio es abductivo y el segundo estadio es deductivo, y circulamos de un estadio al otro hasta que resolvemos el problema o abandonamos. Los componentes del problema de las conexiones son de dos tipos: los componentes estáticos que son los cables y la pila, y los componentes variables que son los interruptores y la bombilla, cada uno de los cuales puede estar en dos estados. Aquí se muestra mediante un diagrama simplificado un estado de un interruptor:



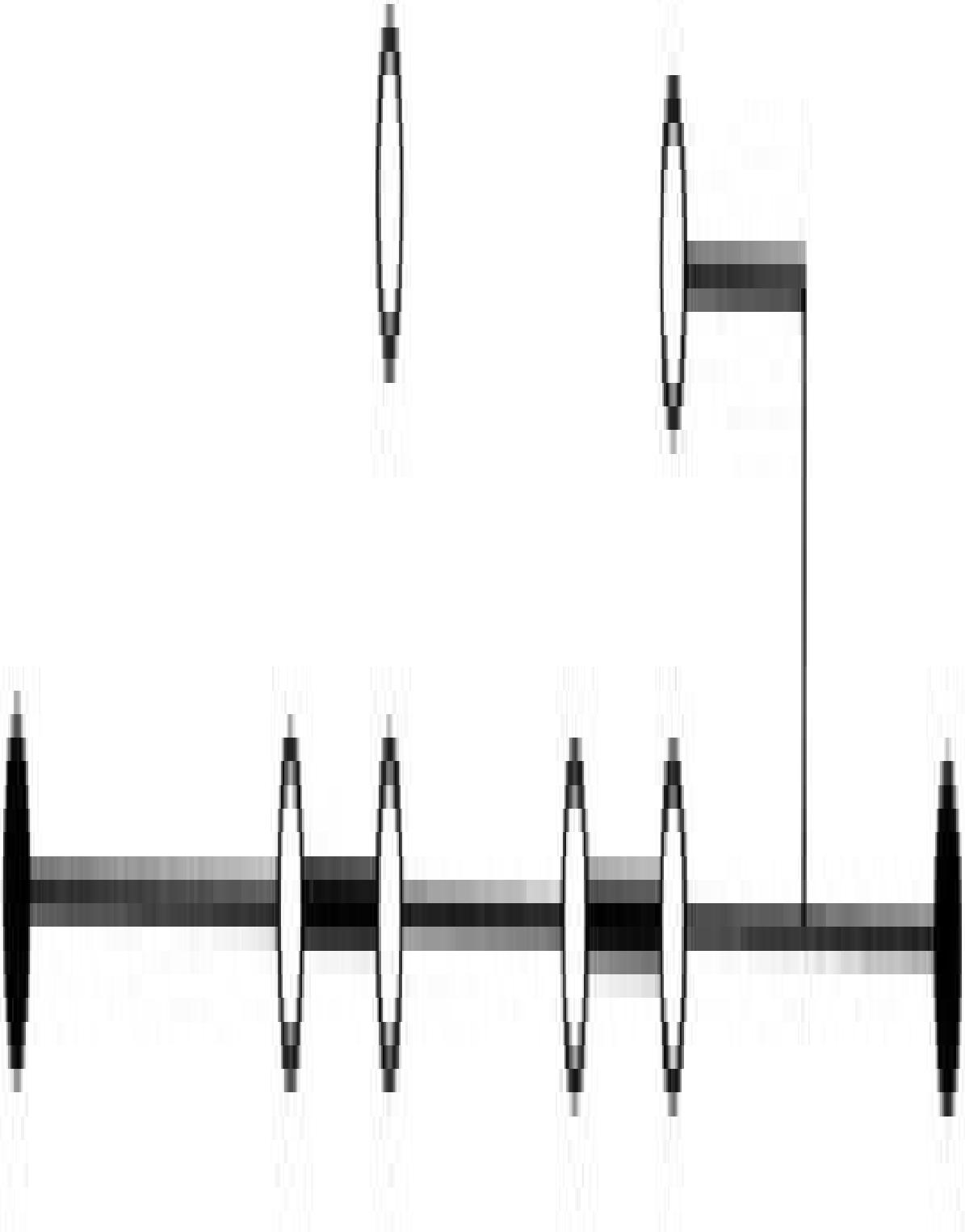
El otro estado es:

Este estado abre el circuito inferior y, si es necesario, forma un circuito para el terminal superior. Los estados de los dos interruptores son la entrada del sistema. El estado de la variable restante, la bombilla, es la salida.

Hay dos tipos principales de estrategia que podemos desarrollar para hacer ingeniería inversa con un sistema. Podemos centrarnos o en las salidas del aparato o en sus entradas. Con las salidas, nos centramos en primer lugar en aquellas que son positivas, es decir, la bombilla se enciende. Una posibilidad para el problema con “o” es que ambos interruptores estén pulsados y la bombilla se encienda. Construimos el circuito más simple posible con dos interruptores:

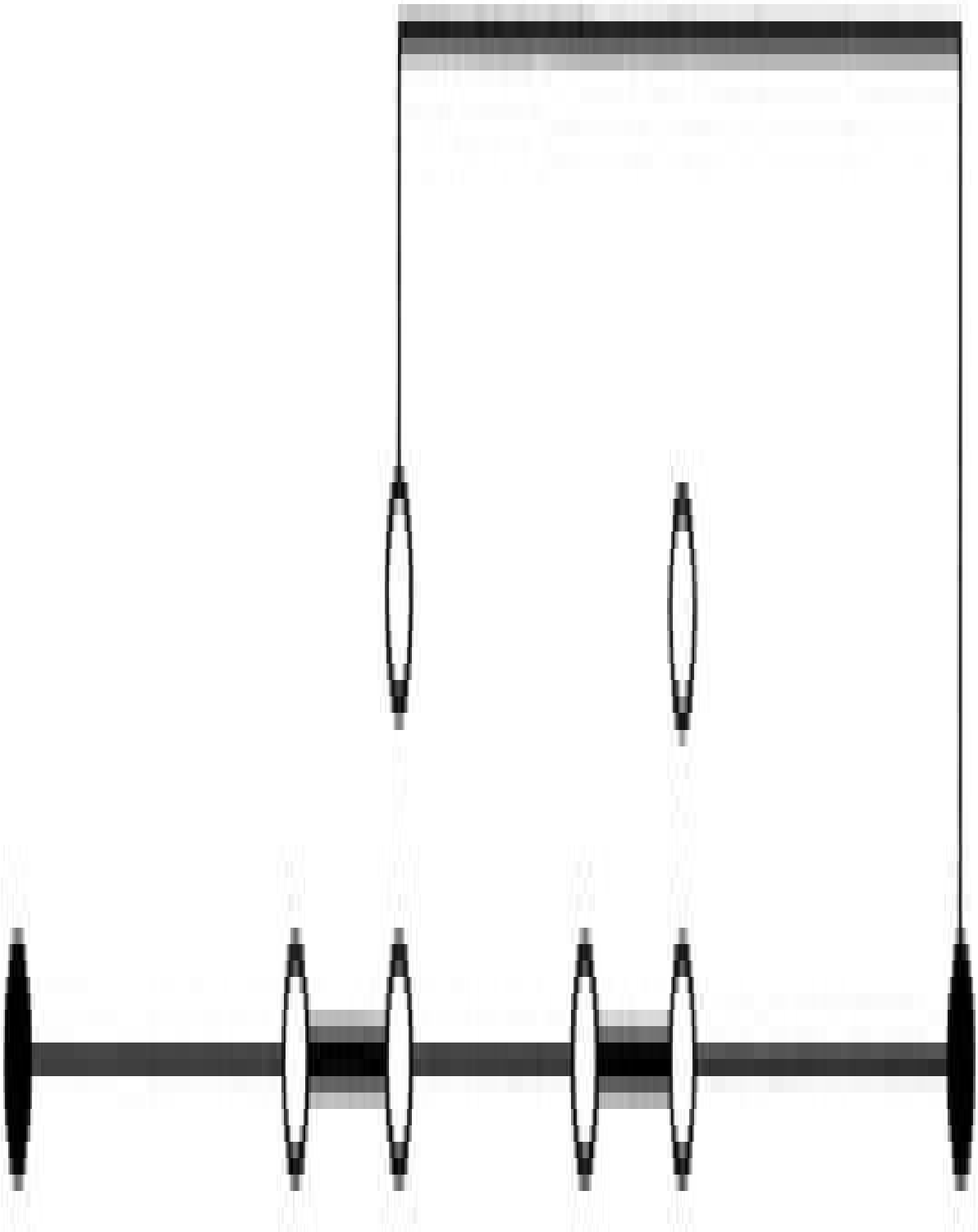


Por sencillez, supongo que en estos diagramas los dos extremos de este cable están conectados a la pila y la bombilla para formar el resto del circuito. Podemos deducir que la bombilla se enciende cuando ambos interruptores están pulsados, y que no se enciende cuando ambos interruptores están sin pulsar. Pero, como también podemos deducir, la bombilla se enciende solo cuando ambos interruptores están conectados. Me refiero a este tipo de circuito como un circuito “y” porque el primer y el segundo interruptor tienen que estar pulsados para que la luz se encienda. Los electricistas dicen que los dos interruptores están conectados “en serie”. Generamos un circuito corregido que incluye una unión en T, como por ejemplo:

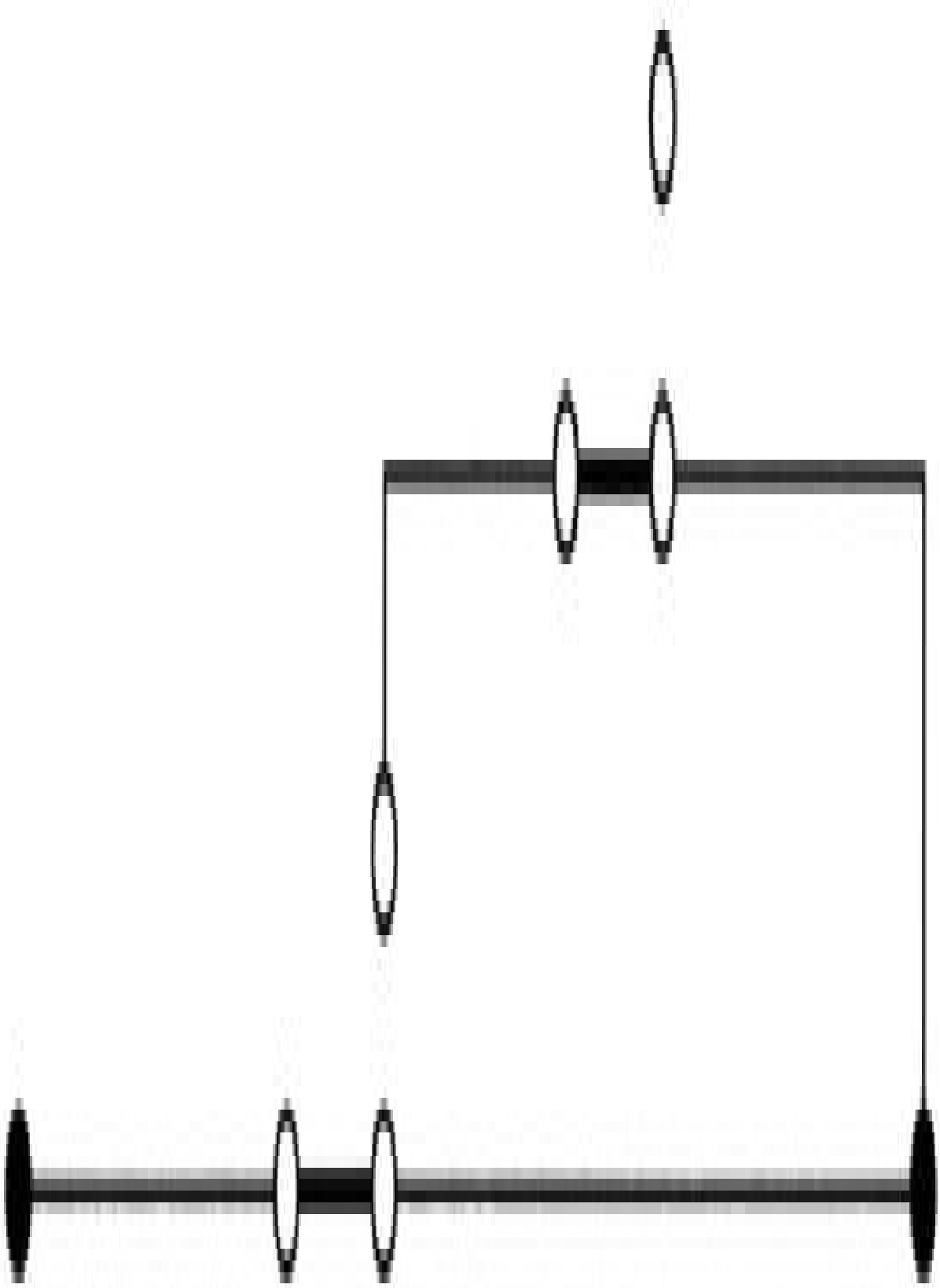


Deducimos que este circuito explica tres de las posibilidades: ambos interruptores pulsados y la bombilla encendida, ambos interruptores sin pulsar y la bombilla apagada y el primer interruptor pulsado y el segundo interruptor sin pulsar y la bombilla encendida. Pero no consigue explicar la cuarta posibilidad en la que el primer interruptor está sin pulsar y el segundo interruptor está pulsado y la bombilla se enciende. Podríamos intentar varias alternativas, como conectar el primer interruptor de la misma forma que el segundo del diagrama anterior. Lamentablemente, la bombilla se enciende independientemente de la posición de los interruptores. Y por tanto vemos que hemos llegado a un callejón sin salida.

Una forma de salir del punto muerto es considerar las salidas en un orden diferente. Ciertamente, podíamos haber empezado con un orden diferente del que he descrito. Empezamos de nuevo con el caso en el que ambos interruptores están pulsados y la bombilla se enciende, y construimos el circuito “y” que mostré anteriormente. Extendemos el circuito para afrontar la posibilidad en la que la bombilla se enciende cuando el primer interruptor está sin pulsar y el segundo interruptor está pulsado:

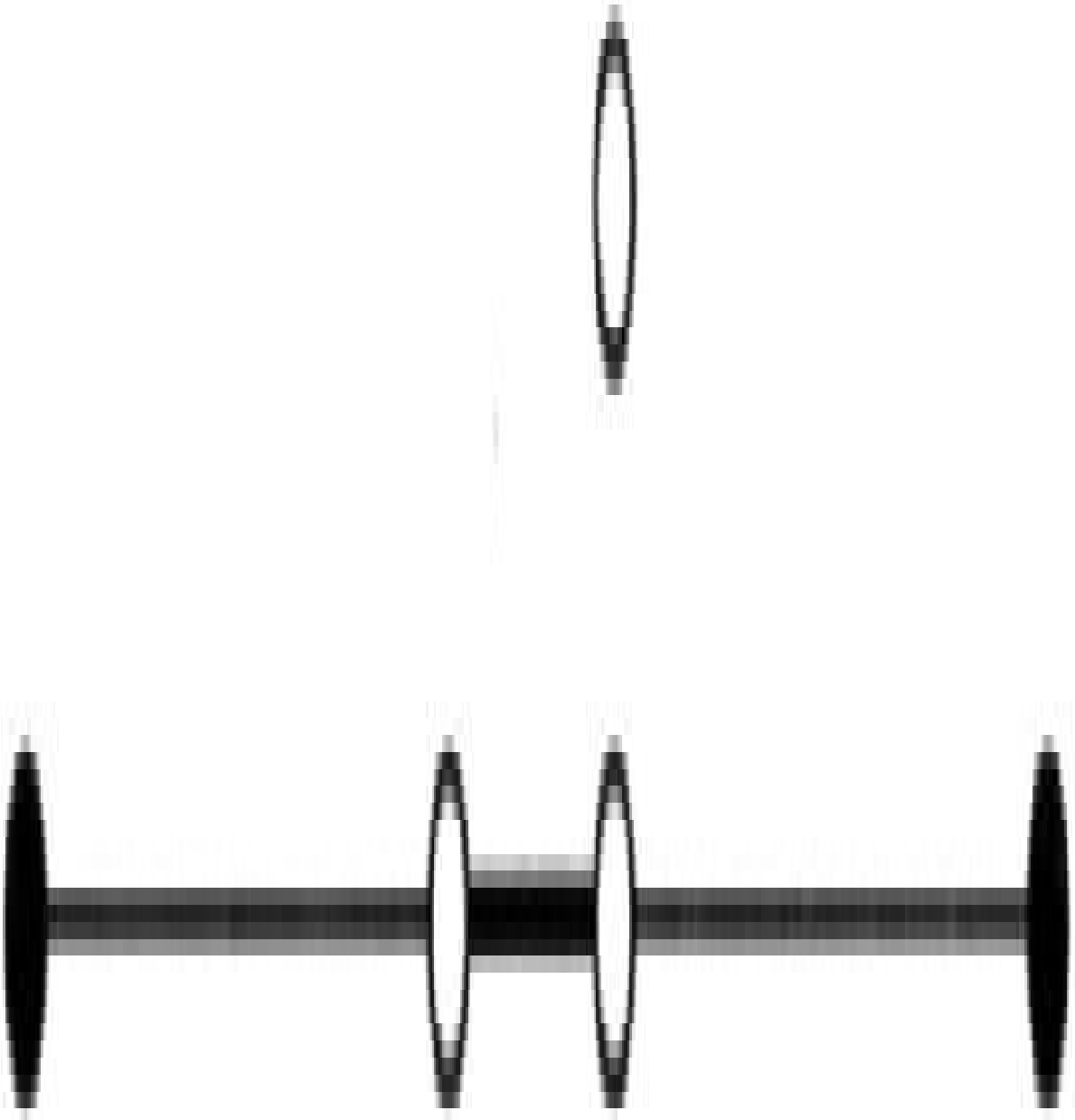


Ahora la bombilla se enciende en tres posibilidades, pero, como podemos deducir, no en las correctas. Podemos intentar trasladar el interruptor al circuito superior:

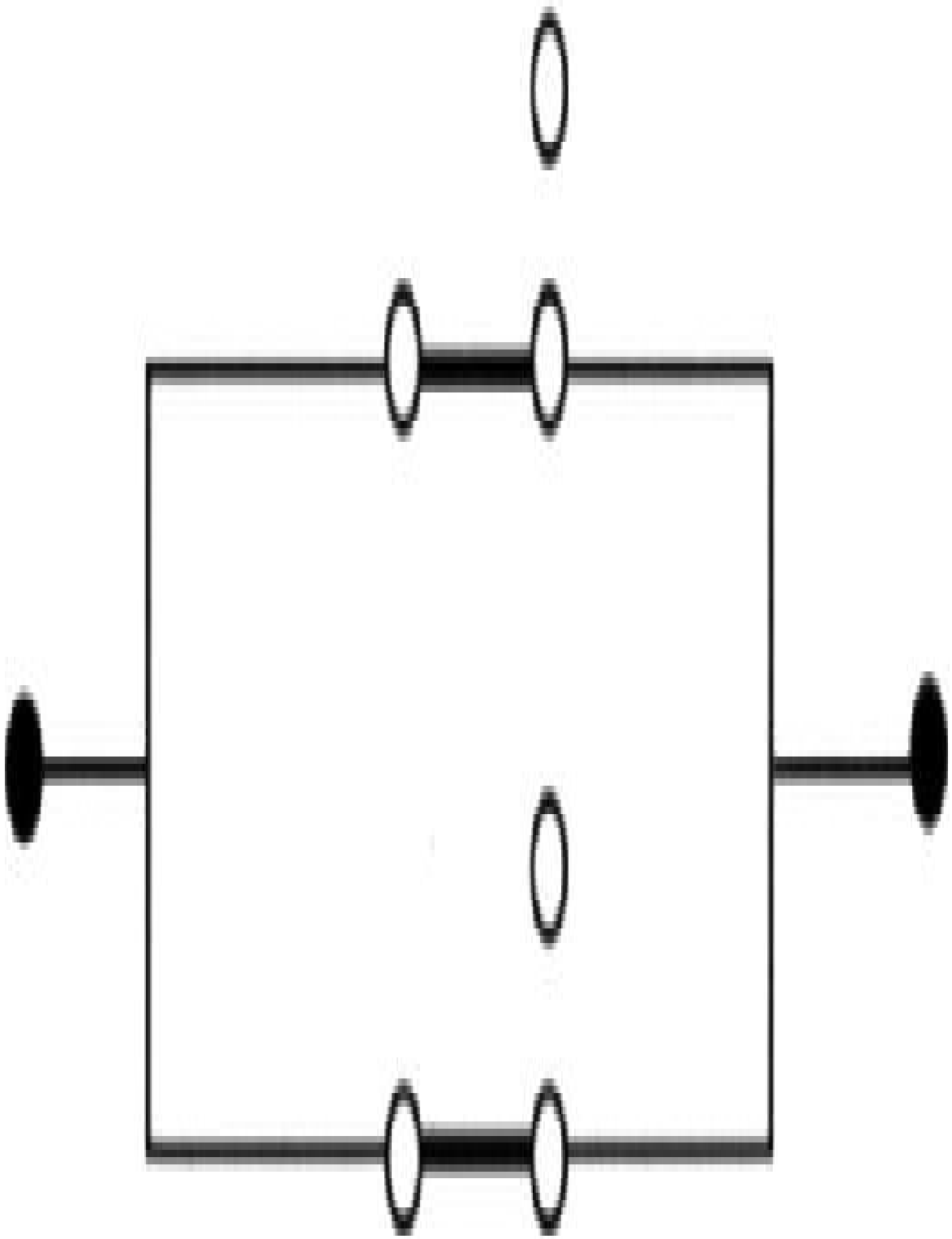


Finalmente, como podemos deducir, la bombilla se enciende cuando ambos interruptores están pulsados, y cuando uno está pulsado y el otro sin pulsar. Y no se enciende cuando ambos interruptores están sin pulsar. Este circuito es una forma de resolver el problema –no la forma estándar de un electricista, y eso tiene la consecuencia de que cuando el primer interruptor está pulsado, el segundo interruptor es irrelevante.

Si en lugar de eso nos centramos en los componentes de la entrada, empezamos con un interruptor y tratamos de conectarlo de forma que su posición dé lugar a las salidas correctas. Este circuito es el punto de partida más simple:



Pero podemos deducir que la bombilla no se puede encender cuando el interruptor está sin pulsar, y por tanto no permite que la bombilla se encienda en algunos casos en los que el interruptor está sin pulsar. En este punto, podemos inferir que cada interruptor enciende la bombilla sin depender del otro, y por tanto necesitamos considerar un circuito independiente para el segundo interruptor:



Este circuito estándar “en paralelo” resuelve el problema: deducimos que la bombilla se enciende siempre que al menos uno de los interruptores esté pulsado.

Las estrategias no dictan en qué orden considerar las posibilidades o cómo enfrentarnos a ellas. Podemos incluso pasar de una estrategia a otra y luego volver atrás. El proceso es no determinista. Pero, cada vez que añadimos un interruptor o un cable al circuito, sea útil el resultado o no, deberíamos deducir al menos una consecuencia de ese paso. Este conocimiento, que se adquiere por deducción, se transfiere de modo que restringe el estadio generativo de abducir un paso. Como con los problemas de formas, se produce un cambio estratégico al utilizar el conocimiento del estadio evaluativo en el estadio generativo.

* * *

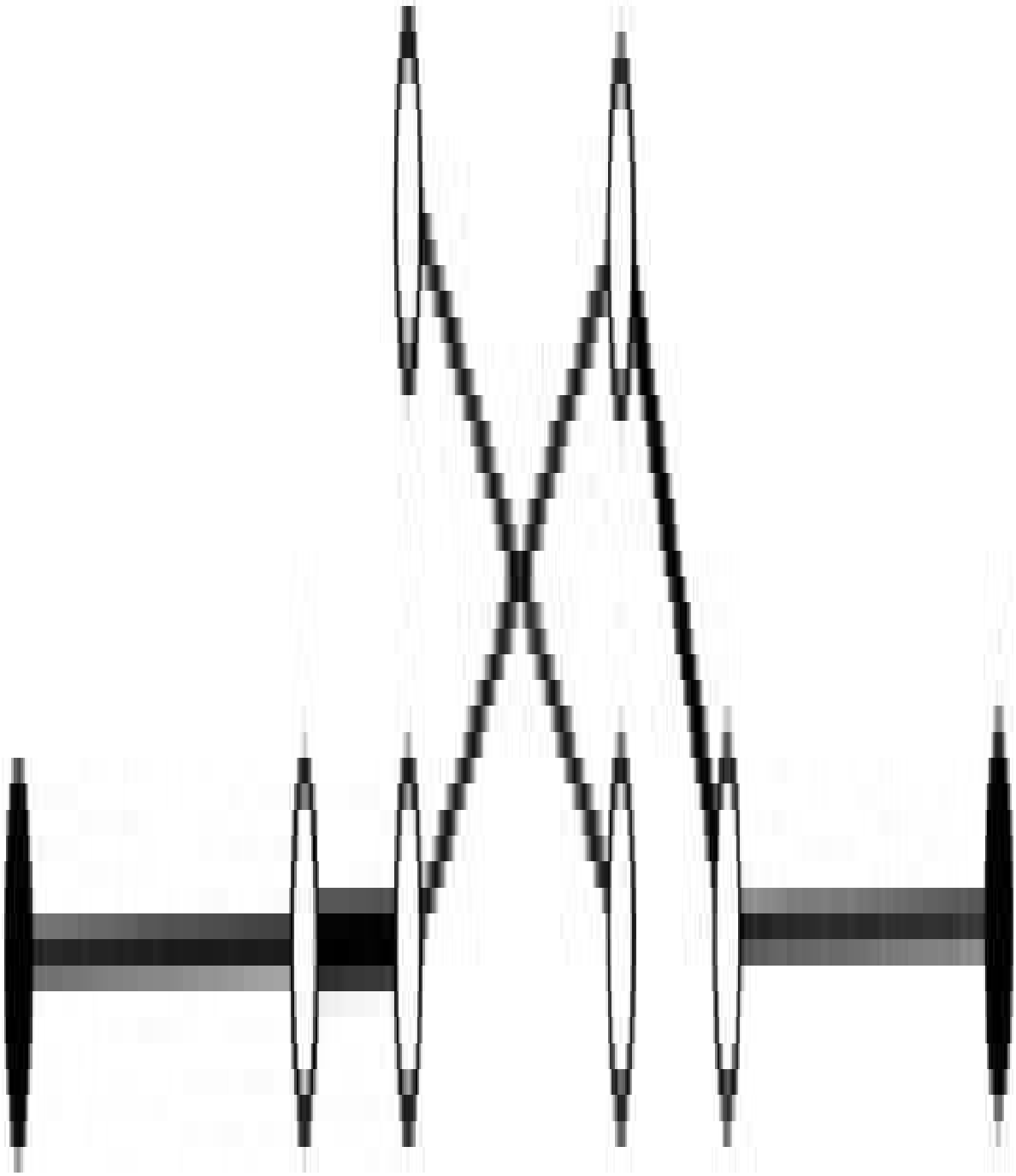
La teoría predice la dificultad de la ingeniería inversa en circuitos booleanos. Una fuente de dificultad es la cantidad de componentes variables, sean componentes de entrada o de salida. Un problema con quince interruptores y veinte bombillas será prácticamente imposible para cualquiera que no sea un experto, a menos que los interruptores tengan que estar conectados en una única serie para que se enciendan todas las bombillas como en una guirnalda navideña. Esas cantidades limitan el número de posibilidades que tenemos que tener en cuenta. Pero el número de salidas positivas en las que la bombilla se enciende es un segundo factor de dificultad en un problema. Un problema con “y” debería ser más fácil que un problema con “o”, porque un problema con “y” tiene una posibilidad positiva, mientras que un problema con “o” tiene tres. Los experimentos han corroborado esta predicción tanto con circuitos eléctricos como con problemas equivalentes en otros dominios, como el control de flujo de agua con grifos. En general, los problemas más fáciles son aquellos en los que hay una sola posibilidad para que el sistema produzca un resultado positivo, y a medida que el número de posibilidades positivas aumenta, sin variar el resto de los elementos, también lo hace la dificultad de la tarea. Con la experiencia, podemos darnos cuenta de que conviene centrarse en los resultados negativos si

son muchos menos que los resultados positivos.

Un tercer factor es sutil, pero crucial. Tiene que ver con hasta qué punto podemos dividir el sistema en subsistemas o componentes independientes. Un componente es independiente cuando su efecto sobre el resultado no se ve afectado por la situación de los otros componentes. Como ejemplo, consideremos de nuevo un circuito con dos interruptores. Usted controla uno de ellos y yo controlo el otro. Si están conectados en un circuito "o", entonces cada uno de ellos puede encender la bombilla sin tener en cuenta el otro. Si usted coloca su interruptor en posición de encendido, la bombilla luce y nada de lo que yo pueda hacer la apaga. En mi caso pasa lo mismo. Sin embargo, para apagar la bombilla ambos debemos apagar nuestros interruptores. Un circuito con los dos interruptores en serie –un circuito "y"– es una imagen especular de esta situación. Usted tiene una forma segura de apagar la bombilla: coloca su interruptor en la posición de apagado y nada de lo que yo pueda hacer enciende la bombilla. En mi caso pasa lo mismo. Sin embargo, para encender la bombilla, ambos debemos pulsar nuestros interruptores.

Ninguno de nosotros puede tener una independencia absoluta. Si usted enciende la bombilla y yo la apago, ¿qué sucedería? Es imposible que la bombilla esté en las dos situaciones, y por tanto uno de nosotros tiene que depender del otro. La independencia dentro de un mismo sistema afecta o a los resultados positivos o a los negativos. Para que cualquiera de nosotros tenga poder absoluto, debemos ser déspotas y controlar sistemas independientes.

La luz en el hall de mi casa se controla con dos interruptores. Un interruptor está al comienzo de la escalera y el otro interruptor al final. Sea cual sea el interruptor que altere, el estado de la luz cambia. ¿No son los dos interruptores independientes uno de otro? No, ni mucho menos. Si usted enciende la luz al comienzo de la escalera, entonces yo puedo apagarla inmediatamente al final, y viceversa. Usted no puede asegurarse de encender la luz más que durante la fracción de segundo que yo tardo en anularle, y viceversa. He aquí un diagrama del circuito:



Como puede usted deducir, la luz se enciende cuando su interruptor está pulsado o bien mi interruptor está pulsado, pero no ambos. El circuito tiene la lógica de una disyunción excluyente. Por tanto, el efecto de la posición de mi interruptor depende de la posición del suyo, y viceversa. Los efectos de los dos interruptores son dependientes uno del otro y no independientes. A diferencia de un problema con “y” u “o”, el circuito no puede descomponerse en componentes independientes. Pero, si un sistema no se puede descomponer en componentes independientes, el que pensemos en los efectos de un componente mientras ignoramos los otros tiene consecuencias. Se deduce que en los sistemas dependientes, como por ejemplo un circuito “o bien”, debería ser más difícil hacer ingeniería inversa que en los sistemas independientes, como un circuito “o”. Con un circuito “o bien” necesitamos pensar en las posiciones de los dos interruptores en conjunto, tanto si nos centramos en los componentes de entrada como en las posibles salidas.

Los experimentos han corroborado esta predicción: los sistemas “o bien” eran más difíciles que los sistemas “y” u “o”. Los participantes inexpertos empleaban más tiempo y tenían menos probabilidad de resolverlos. Por tanto, los factores que afectan a la dificultad de la ingeniería inversa en los sistemas booleanos son tres: el número de componentes variables, el número de posibilidades que dan lugar a resultados positivos y la independencia relativa de los componentes al determinar cualquier tipo de resultado. Una conjetura razonable es que para cualquier campo se mantienen los mismos principios, incluso para la ingeniería inversa de los hornos tostadores.

* * *

Una estrategia natural que solemos adoptar para resolver problemas es trabajar avanzando desde el estado inicial hacia una solución, restringidos por todos los conocimientos que tenemos. Otros tipos de problema inspiran otros tipos de estrategia. Una estrategia famosa, que se remonta a Platón, es empezar con la solución requerida, y trabajar retrocediendo hacia el estado inicial. Esta estrategia se conoce como análisis “medios-fines” porque depende del

conocimiento de los medios de que disponemos para alcanzar nuestros fines. Nuestro cubo tiene un agujero que necesitamos arreglar. El medio, como sabemos, es hacer un tapón de caña. Pero no podemos utilizar la caña porque es demasiado larga. Por tanto, necesitamos un cuchillo. El cuchillo es romo, y por tanto necesitamos afilarlo. Podemos afilarlo en una piedra. Pero la piedra está seca, y por tanto necesitamos mojarla. En este punto, debemos tener cuidado para evitar el problema que surge en la canción folk que Harry Belafonte hizo popular. No deberíamos tratar de ir a buscar el agua con nuestro cubo (tiene un agujero). Deberíamos llevar la piedra al río.

La estrategia medios-fines ha tenido influencia tanto en el estudio sobre cómo resolvemos los problemas como en el desarrollo de programas de ordenador inteligentes. Sin embargo, no es aplicable a los problemas de formas o a los problemas de interruptores. No podemos tener modelos para sus objetivos finales, y trabajar retrocediendo a partir de ellos, porque los objetivos finales son las soluciones de los problemas. Otra estrategia más, que utilizan los matemáticos cuando tratan de comprobar ecuaciones, es un equilibrio entre trabajar avanzando y trabajar retrocediendo. Empiezan a trabajar partiendo de un lado de la ecuación y luego, si se quedan atascados, pasan a trabajar en la dirección opuesta. Un objetivo teórico para los psicólogos es formular una explicación de todas nuestras posibles estrategias para la resolución de problemas. Sin embargo, la dificultad está en tener en cuenta todos los tipos de problema.

* * *

¿Qué pasa con el insight - depende de un proceso mental especial? Para tratar de responder a esta cuestión, propondré una descripción amplia del insight como suceso, y luego me enfrentaré a los procesos que subyacen en él. Empiezo con un problema ilustrativo:

Una forma está constituida por exactamente cuatro cuadrados del mismo tamaño. Quite una pieza para eliminarlos todos sin extremos sueltos.

Si un experimentador plantea este problema a participantes desprevenidos que acaban de resolver un conjunto de problemas de formas del tipo de los anteriores, al principio están perplejos. Es un problema difícil. Puede intentar resolverlo por sí mismo antes de seguir leyendo.

Si no conocía ya la solución al problema, es probable que se haya bloqueado durante un rato o bien que haya abandonado. Pero, si perseveró y resolvió el problema, entonces puede que se haya dado cuenta repentinamente de que un problema de formas podría ser tridimensional. Imaginó entonces una distribución de los cuadrados parecida a la que aparece en la Figura 24.3. Quitó la pieza vertical central, y de ese modo eliminó los cuatro cuadrados, porque se convirtieron en dos rectángulos. El problema es típico de los que los psicólogos han estudiado como ejemplos de insight. Tiene también las características de los problemas de la vida real que giran en torno a un insight.

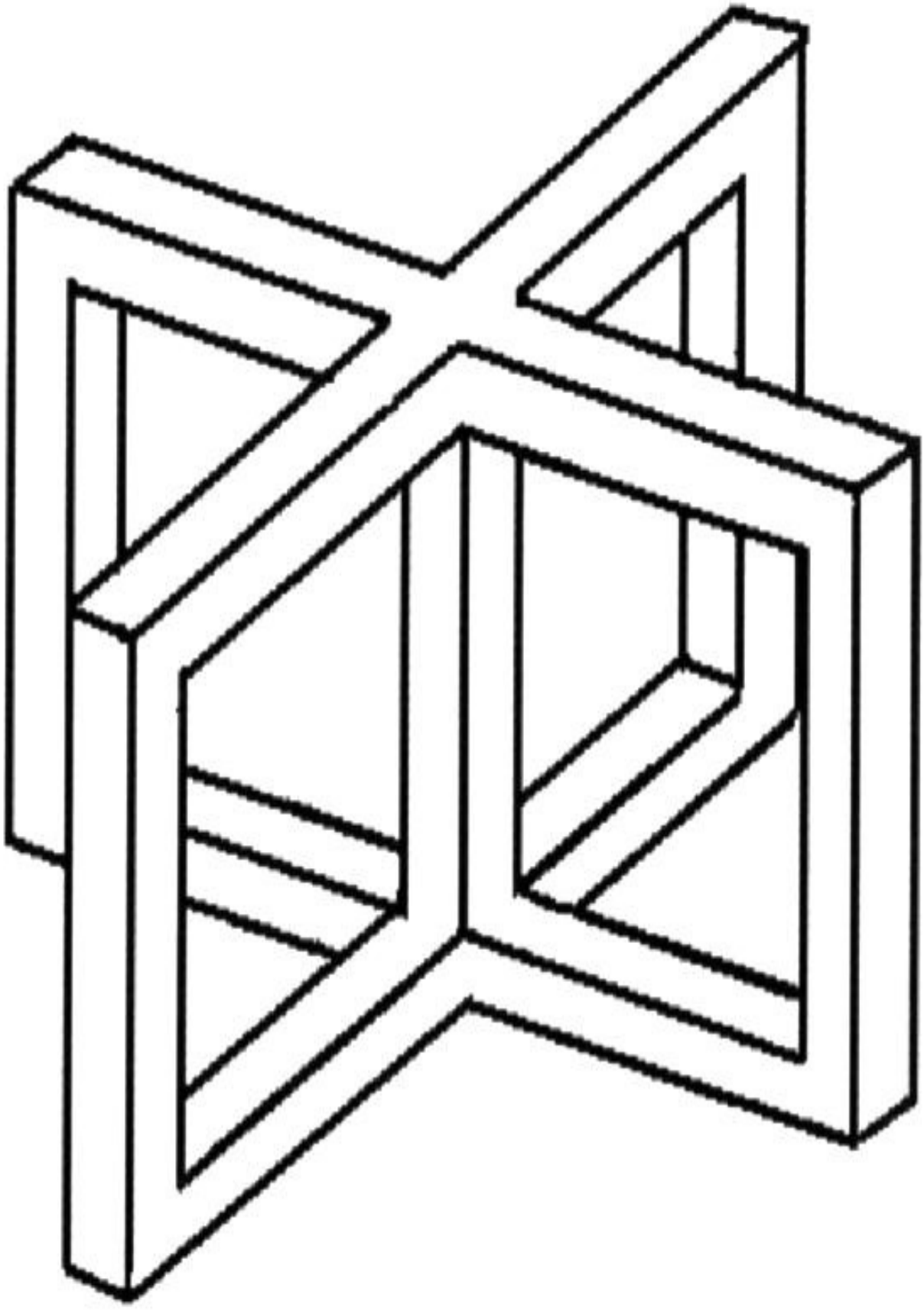


Figura 24.3. Solución tridimensional para un problema de formas.

Cuando tenemos un insight, en un momento determinado estamos en un punto muerto sin ser capaces de resolver un problema, y en el momento siguiente somos conscientes de la solución. Sentimos que se nos ha caído la venda de los ojos, es decir, hubo una discontinuidad en la que se produjo un cambio fundamental en nuestro pensamiento. Si resolvió el problema anterior, es posible que se produjera un cambio en su pensamiento, y puede usted describirlo: estaba pensando en dos dimensiones, y luego pensó en tres dimensiones. No fue consciente del proceso que condujo al cambio, pero fue consciente del cambio y de la solución a la que condujo. Es usted consciente de lo importante que fue el cambio, y puede describírselo a otras personas o darles pistas que pueden capacitarlas para llegar también al insight. Excepto los escépticos más inamovibles, todo el mundo debería aceptar que se producen estos cambios. Su descripción implica que la solución de algunos tipos de problema puede depender del insight, mientras que la de otros tipos de problemas, no. Pero el que se produzca un insight es un asunto subjetivo: depende de si resolvimos el problema, y de si fue como resultado de una discontinuidad en nuestro pensamiento que dio lugar a un cambio significativo del que fuimos conscientes.

La teoría sobre cómo resolvemos los problemas, descrita en este capítulo, contiene un mecanismo que provocaría el insight. La condición previa es que las constricciones que el proceso generativo utiliza en ese momento no consigan resolver el problema. El resultado es un punto muerto. El sistema abductivo registra el punto muerto, y por tanto inicia un intento de modificar las constricciones. El proceso puede comenzar en la conciencia, pero avanza en el estadio generativo inconsciente. Este estadio puede alterar sus constricciones, y las constricciones modificadas pueden dar lugar a una solución. En el caso de un falso insight, pensamos que tenemos una solución, pero, a pesar del cambio en las constricciones, posteriormente deducimos que estábamos equivocados. Un genuino insight es por tanto un fallo inicial al resolver un problema, seguido de un cambio en las constricciones, seguido de una solución. Una vez que se ha producido un insight, puede utilizarse en lo sucesivo para restringir el propio estadio generativo: buscamos inmediatamente una solución tridimensional cuando los intentos bidimensionales fallan. Un cambio en las constricciones está

obligado a ser discontinuo, pero puede estar precedido por cambios inútiles. Un caso especial de cambio en las constricciones da lugar a una nueva representación de un problema, p. ej., un cambio de dos a tres dimensiones, y algunos psicólogos han sostenido que es este tipo de cambio el que resulta crucial para el insight. Ciertamente, otros problemas de formas que encajan en mi descripción de un insight solo dependen de relajar la restricción de que los cuadrados deberían ser del mismo tamaño.

Parte VIII

Razonamiento experto en tecnología, lógica y ciencia

Una hipótesis científica es un marco provisional

para organizar futuras experiencias

Auden

CAPÍTULO 25

Bicicletas voladoras: cómo inventaron el aeroplano los hermanos Wright

A los hermanos Wright les llevó menos de cinco años inventar una nave más pesada que el aire que volara por sí misma y bajo control del piloto. Diseñaron y construyeron su “máquina voladora” y el motor de combustión interna. Sus primeros vuelos se realizaron hace poco más de cien años cerca de Kitty Hawk en los Outer Banks de Carolina del Norte. Habían desarrollado la máquina a partir de una cometa y tres planeadores tripulados. ¿Por qué tuvieron éxito cuando tantos otros no lo consiguieron? Se han hecho muchas conjeturas, que van desde su excepcional habilidad mecánica a la psicodinámica de sus vidas como solteros sin la carga de ninguna relación sexual. Como con cualquier misterio histórico, nunca podremos saberlo con seguridad. Una opinión, por supuesto, es que el genio es inexplicable, y los Wright eran genios. A ellos les habría hecho gracia esta idea[1]. Eran sensatos hombres prácticos de Dayton, Ohio. El propio Wilbur pensaba que no tenían capacidades excepcionales, y que si el tiempo hubiera retrocedido hasta el inicio de sus esfuerzos en 1899, y luego se hubiera puesto en marcha de nuevo, es improbable que la segunda vez hubieran tenido éxito. Solo estaba impresionado por el poco tiempo que les había supuesto. Pero pensaba que habían sido afortunados. Habían sido afortunados, pero no habían sido solo afortunados.

* * *

En esta parte del libro, voy a considerar tres casos de razonamiento experto –en tecnología, lógica y ciencia– con el fin de tratar de entenderlos a la luz de la teoría de modelos. Empiezo por los hermanos Wright. Lo que muestra que fueron extraordinarios es el excepcional adelanto que establecieron respecto a sus rivales. Mucho antes de que ellos hubieran descubierto los principios del

vuelo, otros estaban o bien robando sus ideas o estrellando aviones con fatal regularidad. Podemos eliminar inmediatamente una serie de factores en relación con su éxito –unos porque a los Wright les faltaban y otros porque los tenían sus rivales–. Los Wright no tenían una formación más allá de la Escuela Secundaria, y no tenían preparación científica. No eran ricos. Su negocio de bicicletas les permitía tener tiempo para sus experimentos y para pagar sus máquinas, que no eran muy caras. Calcularon que su primer avión satisfactorio costó menos de 1.000 dólares. Tenían habilidad para la construcción y reparación de máquinas; Orville ya había ideado una prensa de impresión y una calculadora a partir de sus propios diseños. Eran concienzudos, cabezotas y persistentes. Sin embargo, entre su rivales se incluían ingenieros preparados, con habilidad, persistencia y dinero. Samuel Langley, el secretario de la Institución Smithsoniana de Washington, recibió 150.000 dólares del dinero de los impuestos para construir un aeroplano, pero no lo consiguió.

Lo que voy a defender es que los Wright tuvieron éxito porque razonaban mejor que sus rivales. Puede usted decir, en cambio, que los Wright eran más inteligentes que sus rivales. Pero estamos de acuerdo en este caso, porque la inteligencia depende de la capacidad para razonar (véase el Capítulo 1). Los Wright rara vez dijeron nada sobre sus métodos –estaban demasiado ocupados razonando como para tener tiempo de reflexionar sobre cómo estaban razonando–. Además, como he sostenido, los procesos de razonamiento son inconscientes. Solo sus resultados pueden ser conscientes. Mi principal objetivo es mostrar que eran magníficos razonando y extrayendo analogías fructíferas. Cualquier intento de reconstruir el pensamiento de un individuo está obligado a ser especulativo, pero dispongo de pistas repartidas por los escritos de los hermanos y de una teoría del razonamiento.

* * *

El propósito de los Wright era construir un avión, no adquirir conocimiento científico. A finales del siglo diecinueve los científicos miraban con recelo la exploración aeronáutica. El físico Lord Kelvin escribió en 1898 que los esfuerzos en la aviación “solo podrían conducir a la decepción, si se llevaban a cabo con alguna esperanza de que condujeran a una máquina voladora que fuera

útil”. El interés de los hermanos en los principios del vuelo era práctico. En el capítulo anterior, distinguí tres tipos de procesos para el pensamiento creativo. La historia del razonamiento de los Wright deja claro que intentaron reducir al mínimo el “ensayo y error”: no utilizaron una estrategia neodarwinista. Pero ningún inventor empieza conociendo todas las restricciones necesarias para producir el invento satisfactorio. Tener ese conocimiento sería tener ya el invento en mente. Y los Wright tampoco trabajaron de esta forma neolamarckiana. Utilizaron algunas restricciones previas al crear sus diversos aviones, pero también comprobaron sus ideas en la práctica para ver si encontraban otras restricciones, volviendo sobre los resultados para el trabajo posterior, reevaluando esos resultados, etc. Recorrieron muchas veces estos estadios creativos y evaluativos: utilizaron una estrategia de “estadios múltiples”. Como muchos de los que piensan por primera vez en algo, pensaron mucho en qué era lo que querían inventar, y luego utilizaron esos criterios para medir su éxito.

Para evitar el ensayo y error, uno necesita conocimiento, y por tanto si los Wright podían encontrar el conocimiento que necesitaban en la bibliografía, luego sus diseños se ajustaban a los criterios establecidos en ella. Si no podían encontrarlo en la bibliografía, entonces lo desarrollaban por sí mismos. Y, como descubrieron, la bibliografía no siempre tiene razón. Wilbur dio el primer paso para volar en mayo de 1899 cuando escribió a los smithsonianos pidiendo artículos sobre el vuelo y una lista de lecturas recomendadas. Sus fuentes de inspiración inmediatas fueron un libro reciente de ornitología, y la muerte, unos pocos años antes, del pionero alemán de los planeadores Otto Lilienthal, que se había matado estrellándose en una de sus propias máquinas.

Después de tres meses de leer la historia de la aeronáutica, Wilbur se dio cuenta de que teorizar mucho era una tontería. La lección que sacó de la bibliografía era que un aeroplano depende de tres componentes: alas para proporcionarle elevación, un motor para impulsarle y un sistema que permita al piloto controlarlo. Ese análisis lo había hecho Sir George Cayley, el pionero inglés de los planeadores de principios del siglo diecinueve. Pero, ¿qué componente era el más importante?

* * *

Muchos de los contemporáneos de los Wright creían que lo único que faltaba para el éxito era un motor ligero pero potente. En el Capítulo 1 cité a Sir Hiram Maxin en relación con esta cuestión: “Dennos un motor”, dijo, “y muy pronto les daremos una máquina voladora satisfactoria”. Pero dos poderosos motores no habían impedido que su avión se estrellara. Wilbur no puso la prioridad en los motores. Reflejando perfectamente la invención del aeroplano, escribió: “Es posible volar sin motores, pero no sin conocimiento & habilidad”. A posteriori, la afirmación es obvia, pero consideremos el razonamiento de Wilbur:

1. Los motores fallan de vez en cuando. Por tanto, es probable que los motores de aeroplano fallen.

Aquí, hizo una inducción verosímil: los motores sí fallan de vez en cuando, y por tanto es probable que un motor de aeroplano falle. Era posible, pero improbable hasta donde él sabía, que los ingenieros pudieran construir un motor que no fallara nunca.

2. Si el motor falla y el piloto no tiene control sobre el avión, entonces este se estrellará y es probable que el piloto muera.

Había encontrado un claro contraejemplo para la opinión de que un motor resolvería el problema. Puede ser difícil encontrar contraejemplos (véase el Capítulo 16), pero Wilbur tuvo poca dificultad. Uno de sus pocos comentarios sobre la capacidad mental fue: “Es una característica de toda nuestra familia el ser capaces de ver los puntos débiles de cualquier cosa...” Su argumento sobre la prioridad del control era también una inducción a partir del destino de Lilienthal y otros pilotos de planeadores como el inglés Percy Pilcher, que había muerto de la misma forma que Lilienthal. Ambos perdieron el control de sus planeadores. Wilbur infirió por analogía que un avión con un motor que ha fallado es un planeador.

3. Si el motor falla pero el piloto tiene el control del avión, entonces será capaz de aterrizar con seguridad. Por tanto, el control de un avión debe ser anterior al desarrollo de los motores.

La conclusión es una inferencia válida a partir de las premisas (de la 1 a la 3) dada la asunción de que un aterrizaje seguro es preferible a un desastre.

Un rasgo destacable del pensamiento de Wilbur es que pocos pioneros de la aeronáutica pensaron de la misma forma. Querían volar primero y razonar después. Muchos de ellos no sobrevivieron, o bien fueron lo suficientemente inteligentes como para pagar a alguien que hiciera el vuelo por ellos. Uno de los que puso el vuelo por delante del razonamiento fue John J. Montgomery, la primera persona que voló en un planeador en Estados Unidos. El diseño de sus planeadores incluía poco conocimiento de los principios de control. Uno de ellos se soltó desde un globo dejándolo a la deriva a varios cientos de pies, y, después de uno o dos vuelos, quedó fuera de control enviando a su piloto a la muerte. El propio Montgomery murió más tarde, de la misma forma. Wilbur conocía a Montgomery, previó la catástrofe, pero no tuvo forma de evitarla.

La controversia sobre quién fue el primero que inventó el aeroplano no tiene fin. Una razón es que otros contendientes formularon criterios menos restrictivos que los de Wilbur respecto a lo que se debe considerar como un aeroplano. Si el éxito fuera solo cuestión de conseguir una máquina más pesada que el aire que se mantuviera en el aire, entonces podría decirse que otros como Maxin, Langley o Richard Pearse, el pionero neozelandés, habían sido los primeros en volar. Pero un viejo adagio del vuelo es que un buen aterrizaje es aquel del que el piloto puede salir andando, y un gran aterrizaje es aquel que deja la máquina preparada para que otro pueda volar. No hubo buenos aterrizajes, menos aún grandes aterrizajes, hasta los vuelos en Kitty Hawk hace poco más de un siglo. Fueron los primeros en los que el piloto controló la nave.

* * *

Los hermanos establecieron la maniobrabilidad bajo control del piloto como el criterio decisivo para conseguir un aeroplano satisfactorio, y difirieron de otros aviadores en cuanto al tipo de control. Rivales como Maxin y Langley tenían como objetivo una máquina que volara en un equilibrio estable –siguiendo una línea recta sin desviaciones como la que se espera actualmente de un avión de línea–. Los coches, los autobuses, los trenes y los camiones nos llevan de un sitio a otro, y son estables. Se tambalean, pero raras veces pierden el equilibrio. Era natural que los inventores trasladaran esta característica al diseño de aviones. Buscaron la estabilidad, por ejemplo, en una configuración diédrica de las alas de la nave –alas que apuntaban hacia arriba con una ligera forma de V–. Si una ráfaga de aire giraba el avión hacia un lado, entonces suponían que el ala inferior generaría más elevación que el ala superior y, por tanto, causaría el retorno por compensación al vuelo estable. Hubo otros que intentaron desarrollar métodos mecánicos para mantener la estabilidad.

Los Wright ni siquiera tuvieron como objetivo la estabilidad. Wilbur dibujó una analogía a partir de la bicicleta "segura" que tenía las ruedas de igual tamaño. Era una locura reciente. Los hermanos eran ciclistas entusiastas, y su negocio era reparar, vender y fabricar bicicletas. Un bicicleta, sin embargo, no es un vehículo estable. Con la práctica, un conductor aprende a equilibrarla y controlarla. Wilbur sostuvo que la esencia del equilibrio en un avión debería ser también el control más que la estabilidad. Como muchas analogías profundas, la suya se refería no a las características superficiales de la fuente de la analogía –el hecho, pongamos por caso, de que las bicicletas están hechas de metal, o que tienen dos ruedas–, sino a una compleja estructura de relaciones del tipo de las que los psicólogos han postulado que subyacen en las analogías. Los conductores equilibran la bicicleta, y la controlan sobre dos ejes: dirigiéndola a izquierda o derecha, e inclinándola a un lado o al otro. Los pilotos deberían controlar la aeronave de manera similar sobre sus tres ejes. Deberían dirigirla a izquierda o derecha, levantarla de un lado o del otro, y hacerla apuntar hacia arriba o hacia abajo.

Las analogías son herramientas útiles para la abducción, y pueden establecerse entre modelos del mundo (véase el Capítulo 14). La analogía de Wilbur tiene un corolario. Lo mismo que los ciclistas tienen que aprender a mantener el equilibrio de sus máquinas, así también los pilotos tendrían que aprender a controlar su avión. Para los hermanos era crucial desarrollar este dominio, tal como él descubrió. Necesitarían haberlo adquirido mucho antes de añadir un motor. Por tanto, deberían empezar con planeadores.

Las analogías fueron una poderosa herramienta en el pensamiento de los Wright, y a menudo las extrajeron usando como fuente la bicicleta. En muchos aspectos, su máquina voladora era una bicicleta voladora. Los Wright extrajeron también muchas analogías “locales” entre un planeador y otro. Pero la analogía entre el control de las bicicletas y el control del avión es mucho más profunda. Pasa de un dominio a otro, aunque los dos dominios pertenecen a la misma categoría superior de vehículos. Las analogías de este tipo afectan no a propiedades superficiales, sino a asuntos complicados tales como las relaciones causales. Sin embargo, no siempre entendemos el alcance de una analogía. Y lo que es todavía más difícil es encontrar la analogía correcta entre todo el conocimiento que poseemos sobre el mundo. El problema de comparar un problema con todas las analogías que podrían ser útiles no es abarcable. Sin embargo, los Wright estaban en una posición ideal para ver muchas analogías entre las bicicletas y un avión.

* * *

Los pilotos de los primeros planeadores, como Lilienthal, iban colgados de su nave y ejercían el control girando el cuerpo para cambiar el centro de gravedad. Con la vista puesta en el objetivo, Wilbur se dio cuenta de que un avión con motor sería demasiado pesado para ser controlado de esa forma. Podría subirse o bajarse un “timón” horizontal o elevador para controlar la ascensión y el descenso. Pero, ¿cómo podría controlarse la inclinación o el giro? En el verano de 1899, Wilbur tuvo un insight clave a partir de una analogía con los pájaros. Escribió sobre los buitres: “Si el borde trasero de la punta del ala derecha se gira hacia arriba y el de la izquierda hacia abajo el pájaro se convierte en un molino de viento en marcha e instantáneamente empieza a girar, siendo el eje una línea que va desde la cabeza a la cola... En el aparato que intento emplear hago uso del principio de torsión.” Su idea era girar las alas izquierda y derecha de la nave en direcciones opuestas de forma que el eje director de una de ellas formara un mayor ángulo que el de la otra respecto al aire que les llegaba. Como resultado de esta “deformación” de las alas, las dos alas generarían una elevación diferente, y por tanto la máquina se escoraría y giraría. Aquí, de nuevo, reaparece la bicicleta. Algunos de los primeros aviadores habían asumido que el avión giraría en un plano horizontal como los automóviles. Los hermanos se

dieron cuenta de que la bicicleta era mejor como analogía. Para girar, un ciclista debe inclinar una bicicleta: un piloto debe ladear un avión. Orville, que para entonces se había unido al proyecto de Wilbur, sugirió que los extremos exteriores de las alas podrían pivotar sobre ejes metálicos, que tendrían engranajes para moverse en direcciones opuestas cuando el piloto tirara de una palanca. Este sistema anticipaba los modernos alerones, pero los ejes de metal y los engranajes habrían sido demasiado pesados para el planeador de los Wright.

El siguiente paso en el pensamiento de Wilbur se basó en una analogía que describí anteriormente (en el Capítulo 14). Para que las alas se deformen en direcciones opuestas, deben ser flexibles. Pero, si las alas son flexibles, sus extremos podrían oscilar arriba y abajo provocando algo peor que la vergüenza. La solución a este problema llegó cuando Wilbur observó una analogía entre la caja de una cámara de bicicleta que retorció por casualidad y las alas de un biplano. Más tarde escribió, “Si haces un tubo de cartón cuadrado de dos pulgadas de diámetro y ocho o diez de largo y eliges dos lados para representar los planos, verás inmediatamente el efecto de torsión que produce el mover un extremo del plano superior hacia adelante y el otro hacia atrás, y cómo este efecto se alcanza sin sacrificar la rigidez lateral”. La solución era construir un biplano. Sus alas podrían girarse al completo para producir una deformación opuesta de las alas de la derecha y de la izquierda sin ningún riesgo de que oscilaran. La suerte favorece a la mente preparada, como se supone que dijo Pasteur. Wilbur estaba preparado para este insight, porque había estado pensando sobre cómo deformar las alas de un avión. La cuestión dependía de visualizar los lados superior e inferior de la caja retorcida como si fuera un biplano. Tenía una capacidad inusual para visualizar soluciones para los problemas. Él mismo escribió: “Mi imaginación dibuja las cosas más vívidamente que mis ojos.” Se han escrito programas de ordenador para extraer analogías, pero, como mencioné en un capítulo anterior al discutir este ejemplo, representan el conocimiento de una forma lingüística. Sin embargo, los expertos en hacer analogías en vez de representaciones lingüísticas utilizan modelos del mundo.

Los hermanos construyeron con bambú y papel de seda un modelo real de las alas. El diseño parecía viable, y por tanto en julio de 1899 se pusieron a trabajar en la construcción de una cometa biplano con el mismo diseño y una envergadura de cinco pies. El operador controlaba la cometa con dos varillas, una en cada mano, de las que salían dos cuerdas, una al ala superior y la otra al ala inferior, de forma que los extremos de las alas pudieran girarse en direcciones opuestas. Wilbur probó la cometa y confirmó que el ala deformada

funcionaba. Ladeaba la cometa de un lado al otro exactamente como él había imaginado.

Los Wright concibieron con los mismos principios un planeador capaz de transportar a un hombre, y volvieron al diseño de las alas. Las alas de un avión producen elevación: el aire viaja a la misma velocidad por encima y por debajo del ala. El recorrido que hace sobre la parte superior de la superficie combada del ala es más largo que el que hace por debajo de ella, y por tanto la presión del aire es menor encima del ala que debajo. Esta diferencia de presión eleva el ala. Por supuesto, el asunto es más complicado, porque los aviones pueden volar boca abajo: lo que importa también es el ángulo del ala respecto al aire que le llega. Como hace el cuerpo del avión, el ala también ofrece resistencia al paso del aire. Esta resistencia se conoce como resistencia al avance [drag]. Toda ala que se mueve a través del aire produce ambas cosas: elevación, fuerza perpendicular respecto al flujo de aire y resistencia al avance, fuerza que empuja el ala hacia atrás, y por tanto en la misma dirección que el flujo de aire. Las intensidades de estas fuerzas varían en función de la forma de la sección del ala y del ángulo que esta forma con el aire. Lilienthal había publicado una tabla con estos datos para las alas de sus planeadores, y los hermanos utilizaron esta tabla en su diseño. Era un biplano con un “timón” horizontal sobresaliendo por la parte delantera, que el piloto podía flexionar hacia arriba y hacia abajo para controlar la subida o el descenso. Debería levantar a un hombre en el aire con un viento moderado.

* * *

En aquella época, los planeadores se lanzaban desde la cima de una colina o se dejaban caer desde un globo de forma que la gravedad aceleraba la nave por el aire. Wilbur sabía que este procedimiento era peligroso porque la nave estaba enseguida a mucha altura sobre el suelo. Lilienthal había muerto en un choque de ese tipo. Pero, ¿era un lanzamiento desde lo alto de una colina el único método para poner en el aire un planeador?

Wilbur resolvió el problema utilizando el método de cambiar el “fondo” por la “figura”. Cuando examinamos un dibujo de un objeto vemos el objeto –la

figura– con un fondo detrás. Pero con un leve esfuerzo podemos cambiar el fondo por la figura y centrarnos en el fondo y la forma que la figura crea en él. Para planear, el aire se tiene que mover sobre el ala del planeador lo suficientemente rápido como para producir una elevación. Wilbur razonó que había dos posibilidades: o bien el planeador podría acelerarse a través del aire –el foco en la figura–, o bien el aire podría acelerarse respecto al planeador –el foco en el fondo–. La primera posibilidad podría contar con la gravedad, pero eso era peligroso. Sería mejor utilizar una catapulta para lanzar el planeador a través del aire. Eso sería más seguro, pero el diseño de una catapulta eficaz sería un reto. La segunda posibilidad era acelerar el aire respecto al planeador. Eso también sería más seguro, porque el planeador podría mantenerse más cerca del suelo, y quizá también podría atarse a una torre con un contrapeso que evitara su caída. Pero, ¿cómo podría acelerarse el aire respecto a las alas?

Podríamos pensar en un gran propulsor o algún tipo de túnel de viento. Wilbur tuvo una idea más simple. Consultó las tablas de la Oficina Nacional de Meteorología sobre las velocidades medias de los vientos en diferentes lugares. Eligió Kitty Hawk en los Outer Banks de Carolina del Norte, una larga y aislada franja de playa separada de la costa, que se extendía entre un estrecho y el océano, porque los vientos dominantes tenían una media de quince a veinte millas por hora. Después de un año muy atareado en la tienda de bicicletas, Wilbur llegó a Kitty Hawk en septiembre de 1900 tras un largo viaje, que incluyó una peligrosa travesía en barco a través del estrecho en una goleta podrida. Había comprado los palos para las alas del planeador en una tienda local, y lo más largo que pudo encontrar era dos pies [unos 70 cm] más corto que las especificaciones de su diseño. Llegaron el resto de los componentes, así como Orville, y los hermanos montaron el planeador, ajustando la tela para que encajara en los palos más cortos. En octubre hicieron volar el planeador, primero como una cometa, luego con Wilbur pilotando un planeador atado y luego pilotando en vuelos libres. Iba tumbado en el ala inferior para manejar los controles. Parecía peligroso, pero los hermanos descubrieron enseguida que era seguro. Tenía la ventaja de reducir la resistencia al avance en comparación con la que supone que el piloto vaya sentado. Wilbur empleó un tiempo total de unos dos minutos en vuelo libre. Como mencioné, pensaba que era vital para los inventores aprender cómo hacer volar su nave a medida que la desarrollaban (sin matarse). Durante mucho tiempo, él fue el único piloto. Y, en bien de la seguridad, los hermanos volaron juntos solo una vez, muchos años más tarde después de que Orville hubiera acompañado a su padre en su primer vuelo en una máquina voladora con motor.

Según la tabla de Lilienthal, el planeador debería haber elevado a Wilbur con un viento ligeramente superior a veinte millas por hora con las alas formando un ángulo de unos tres grados respecto a él. En realidad, el viento tuvo que superar las veinticinco millas por hora y las alas formar un ángulo de casi veinte grados. A partir de sus conocimientos de aeronáutica, Wilbur dedujo que había dos posibilidades que podían explicar la escasa elevación: el hecho de que los hermanos hubieran utilizado un corte transversal menos peraltado que el del ala de Lilienthal o que habían utilizado alas más pequeñas que las de su diseño original. Hacia finales de octubre, dejaron el planeador en Kitty Hawk para que los habitantes locales lo desguazaran –el tejido de satén se utilizó para hacer vestidos para dos jóvenes– y regresaron a Dayton.

* * *

Volvieron a Kitty Hawk a principios de julio de 1901. Habían diseñado un segundo planeador, el cual tenía en cuenta los vientos dominantes en Kitty Hawk. Con una envergadura de veintidós pies [casi siete metros], era el planeador más grande jamás construido. Se arriesgaron haciendo dos cambios no ensayados en las alas. Introdujeron un borde anterior romo, y volvieron al peralte de Lilienthal. El planeador fue una completa desilusión. El vuelo era errático, la elevación era un tercio de la prevista, y el control era impredecible. Con frecuencia caía en picado hasta el suelo o bien, cuando Wilbur trataba de corregir la caída, entraba en pérdida. Una pérdida de sustentación se produce cuando el ángulo de las alas de un avión respecto al flujo de aire se hace demasiado grande, se establece una turbulencia sobre ellas, pierden toda elevación y la nave pasa a tener las condiciones para el vuelo de una verdadera bicicleta. A diferencia de casi todos los aviones modernos, el avión de los Wright tenía un diseño en el que el elevador –el timón horizontal que controla la subida y bajada– sobresalía delante de las alas como un cuello de pato. Esta configuración se conoce todavía en Francia como el diseño “canard”. Se muestra en la Figura 25.1, que se basa en el dibujo que los Wright enviaron en 1903 con su solicitud de patente para el alabeo del ala. Posteriormente los hermanos utilizaron esta misma configuración en su nave con motor. Descubrieron que tenía una ventaja imprevista. Cuando una nave entra en pérdida, no sufre una caída letal en picado contra el suelo como el planeador de Lilienthal, sino que cae plana como un

paracaídas. El canard salvó el cuello de su piloto varias veces.

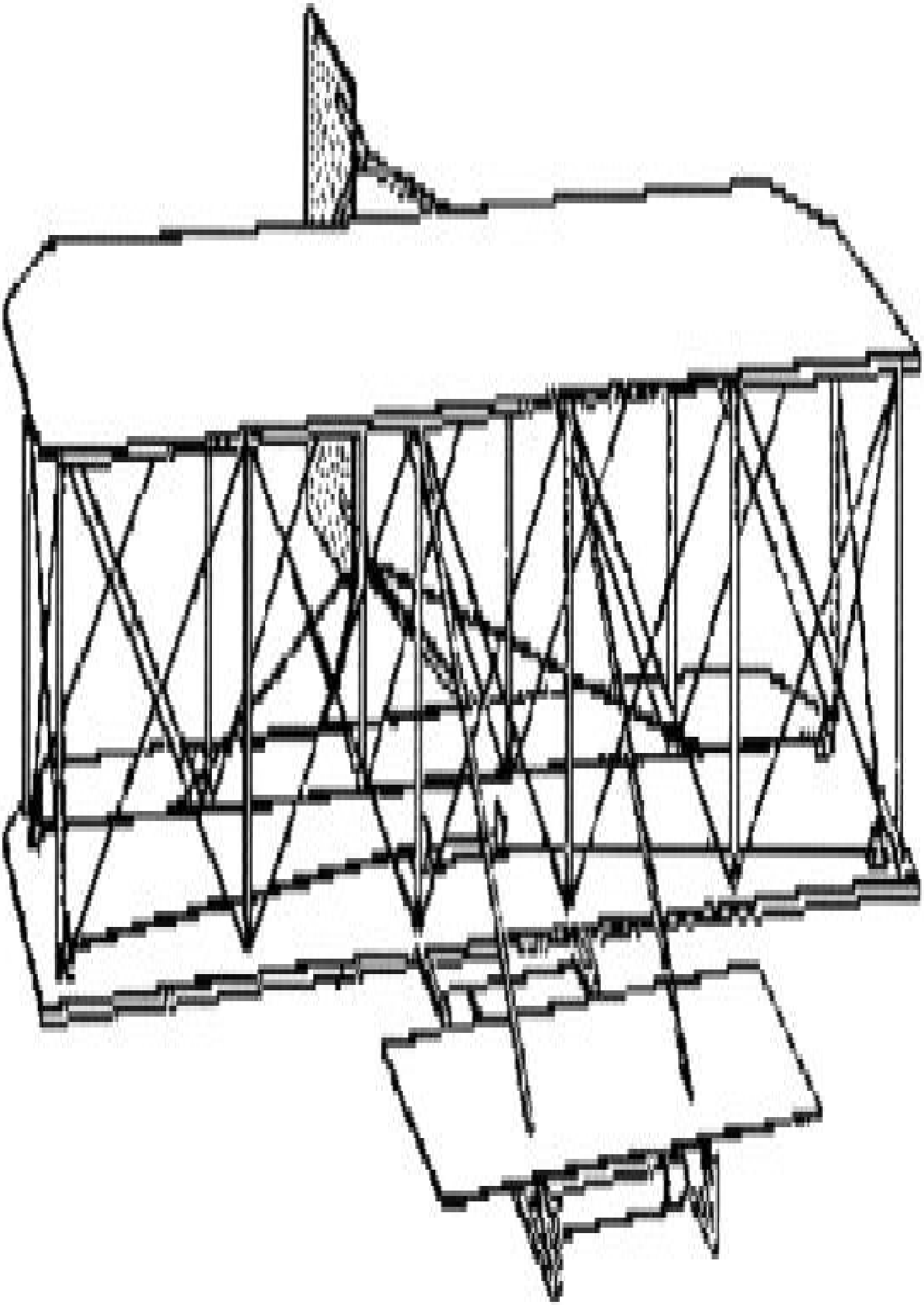


Fig. 25.1. Diagrama basado en la solicitud de patente de los Wrigth de 1903. Muestra el diseño genérico de su nave con el elevador horizontal en la parte delantera y el timón vertical en la parte de atrás.

En el centro de los problemas del segundo planeador estaba la dificultad para controlar el centro de presión del aire. La presión del aire bajo las alas varía de una parte del ala a otra, y tiene un centro en torno al cual el resto de las presiones tienen una distribución similar. Para un equilibrio perfecto, el centro de presión del aire tiene que coincidir con el centro de gravedad del avión. Pero, como dijo Wilbur, los dos tenían “una casi ilimitada incompatibilidad de caracteres que impide que se mantengan juntos pacíficamente durante un solo instante, de manera que el operador, que en este caso actúa como pacificador, a menudo resulta dañado mientras intenta reconciliarles”. El diseño de las alas, pensaban, debería haber minimizado estos problemas, pero quizá había sido un error volver al peralte de Lilienthal.

Los Wright detuvieron los vuelos del planeador para comprobar qué estaba pasando con el centro de presión del aire. Pero, ¿cómo? Puede pararse a pensar por un momento qué haría usted, y enfrentar su capacidad a la ingenuidad de los hermanos. El problema es determinar dónde se sitúa el centro de presión del aire en relación con el centro de gravedad, teniendo solo unos pocos materiales y herramientas rudimentarias, estando a millas de cualquier sitio, y cuando un fuerte viento le azota sin pausa. Nadie de aquellos a quienes he planteado este problema lo ha resuelto hasta ahora. La solución se resume en la Tabla 25.1.

Tabla 25.1. Cómo determinaron los Wright el movimiento del centro de presión del aire en las alas de su segundo planeador

1. Desmantelaron el planeador y volaron el ala superior sola como una cometa utilizando dos cuerdas atadas a su borde delantero.

2. La volaron con un viento ligero y observaron que la parte delantera del ala se torcía hacia arriba, tirando de las cuerdas casi en vertical en relación con sus manos. Infirieron que el centro de presión se había desplazado delante del centro de gravedad del ala, empujando hacia arriba su borde frontal.

3. La volaron con un aire más frío y observaron que el ángulo del ala respecto al viento era menor y las cuerdas flotaban casi horizontalmente entre sus manos y el borde frontal del ala. Infirieron que el centro de presión coincidía con el centro de gravedad, y el ala estaba en equilibrio.

4. La volaron con un viento más fuerte, y observaron que la parte delantera del ala caía y tiraba de las cuerdas hacia abajo entre sus manos y el borde frontal. Infirieron que el centro de presión se había desplazado detrás del centro de gravedad, empujando hacia arriba la parte trasera del ala. Este último movimiento del centro de presión hacia la parte trasera del ala sería lo que empujaría repentinamente su planeador hacia abajo.

5. Infirieron que el gran peralte del diseño de Lilienthal provocaba estos desplazamientos del centro de presión del aire.

Los hermanos rehicieron las alas para cambiar el peralte a uno más estrecho, y construyeron un borde anterior más afilado. Las modificaciones resolvieron el problema del control. Wilbur podía ahora pasar casi rozando el suelo utilizando el elevador para seguir sus ondulaciones. Intentaron entonces el alabeo del ala para controlar la inclinación. No funcionó. Cuando Wilbur alabeaba las alas de los biplanos para ladearse a la izquierda, las alas de la derecha subían y las alas de la izquierda bajaban. Pero la maniobra tenía una extraña consecuencia. Las alas inferiores se movían ahora en el aire más deprisa que las alas superiores, y por tanto la nave viraba a la derecha aunque estuviera inclinada hacia la

izquierda. Este efecto desbarataba tanto el avión como la teoría de los hermanos sobre el alabeo del ala.

Volvieron a Dayton abatidos. La elevación del planeador era demasiado baja, el alabeo del ala había fallado. Se supone que Wilbur dijo en el tren de vuelta a casa: “ni en cien años volaría el hombre”.

* * *

El planeador no había conseguido producir la elevación predicha. Los hermanos habían encontrado una discrepancia entre las consecuencias de sus creencias y la incontrovertible realidad. Tenía que razonar en sentido inverso a partir de la inconsistencia hasta llegar a una explicación consistente de lo que estaba pasando (véase el Capítulo 23). Wilbur infirió tres posibles causas para la escasa elevación: los hermanos habían cambiado las alas a un peralte más estrecho, la tela de las alas podía no haber sido suficientemente tupida o los datos de la tabla de Lilienthal que predecían la elevación podían haber estado equivocados.

En septiembre de 1901, Wilbur dio su primera conferencia pública. Era el portavoz oficial de los hermanos y durante muchos años el elegido como escritor; Orville no habló en público en toda su vida. La conferencia de Wilbur fue sobre “Algunos experimentos aeronáuticos” y la dio en la Sociedad Occidental de Ingenieros [Western Society of Engineers] en Chicago. Su preparación le llevó a cavilar sobre el problema de la elevación. Los hermanos querían comprobar la tabla de Lilienthal, que estaba basada en las medidas que había hecho utilizando un brazo giratorio con una pequeña sección del ala en un extremo y un contrapeso en el otro. El método podría no haber sido fiable.

Los hermanos necesitaban una forma de medir la elevación producida por diversos tipos de viento. Su siguiente paso se basó en su capacidad para imaginar una herramienta mecánica que pudiera producir la respuesta a un misterio teórico. Abandonaron la gravedad como fuerza compensatoria, y utilizaron en su lugar la resistencia del aire. Imaginaron una pequeña superficie plana enfrentada en ángulo recto con el viento. No produciría elevación de ningún tipo, sino pura resistencia. La compararían con un ala vertical peraltada cuyo borde formara diferentes ángulos con el viento. La tabla de Lilienthal predeciría con qué ángulo

respecto al viento la elevación (y la resistencia) creadas por el ala deberían equilibrar la resistencia creada por el plato plano. Montaron el plato plano y el ala sobre una rueda de bicicleta horizontal. El plato y el ala se mantenían verticales y la distancia entre ambos era un cuarto de la circunferencia de la rueda. Si el viento fuera constante, la herramienta respondería: cuando las fuerzas se equilibraran la rueda se mantendría estacionaria; cuando no se equilibraran, giraría. Sin embargo, los vientos en Dayton no eran constantes. Por tanto, montaron la rueda horizontalmente sobre un eje en la parte delantera del manillar de una bicicleta, y sorprendieron a los habitantes locales pedaleando con el artilugio por la ciudad a una velocidad uniforme, tanto a favor del viento como en su contra. Descubrieron que, para equilibrar la resistencia en el plato plano, el ángulo del ala con el viento tenía que ser mucho mayor que la predicción de Lilienthal. Sus datos parecían ser erróneos, pero podía ser que el método que ellos habían utilizado no fuera más fiable que el ala giratoria de Lilienthal.

Construyeron un pequeño túnel de viento, y luego uno más grande, tomándose un mes para asegurar que el aire fluía en una dirección constante y uniforme. Dentro del túnel de viento, insertaron un nuevo sistema de equilibrado que comparaba de nuevo un ala con una superficie plana. Una única superficie plana creaba demasiada turbulencia, y por tanto en su lugar utilizaron cuatro tiras estrechas. El ala estaba montada sobre esas tiras y conectada con ellas mediante un sistema de radios de bicicleta de forma que la elevación del ala y la resistencia podían compararse con la resistencia de las tiras. Después de cada lectura inicial, reajustaban el equilibrado para corregir la resistencia del ala, y medían solo su elevación. Un segundo equilibrado medía la relación entre la elevación y la resistencia para permitirles calcular la resistencia del ala.

Los túneles de viento ya no son necesarios en el diseño de aviones. En su lugar, programas de ordenador simulan el flujo de aire sobre las superficies de un aeroplano. Los Wright tenían una capacidad similar para imaginar el flujo de aire sobre las superficies. Esta capacidad les bastó para construir sistemas de equilibrado en túneles de viento que dieran lugar a medidas exactas a partir de las cuales pudieran valorar la elevación y la resistencia. Llevaron a cabo una serie de pruebas sistemáticas utilizando diferentes formas para la sección transversal del ala, las cuales formaban diversos ángulos respecto al flujo de aire. Estas observaciones confirmaron el hecho de que los datos de Lilienthal eran erróneos. Los hermanos descubrieron también que un ala estrecha y larga daba lugar a una elevación mayor que un ala corta y ancha de la misma superficie.

Otros habían construido antes túneles de viento, pero ellos fueron los primeros en utilizarlos para hacer medidas exactas. Cuando el tiempo que ellos habían establecido para realizar estos experimentos se agotó, los dejaron, y volvieron a sus asuntos.

Los hermanos diseñaron su tercer planeador, teniendo en cuenta sus resultados en el túnel de viento. Tenían que corregir también el problema con el alabeo del ala. Inclinaba el planeador hacia un lado, pero entonces las alas inferiores adelantaban a las superiores y la nave giraba en redondo en la dirección errónea. Necesitaban una forma de desacelerar las alas inferiores. Una vez más, imaginaron el flujo de aire sobre su planeador. Alguna superficie de la nave tendría que actuar cortando el aire en la parte que estuviera más baja cuando la nave se inclinaba. Imaginaron dos colas verticales fijas detrás de las alas. Cuando el planeador se inclinara hacia un lado, las colas cortarían el viento para frenar a las alas inferiores.

A finales de agosto de 1902, viajaron una vez más a Kitty Hawk. Montaron su tercer planeador, y lo probaron, como de costumbre, como una cometa y luego en planeos cortos. En su primer viaje a la zona, Wilbur había registrado sus observaciones sobre los pájaros. Anotó, por ejemplo, que el buitre con sus alas en un ángulo diédrico tenía una dificultad mayor para mantener el equilibrio con vientos fuertes que las águilas y halcones que mantenían sus alas equilibradas. El ángulo diédrico aumentaba la perturbación producida por las ráfagas de viento laterales. Los hermanos habían equipado su primer planeador con un ángulo diédrico, y lo encontraron insatisfactorio. Rehicieron las alas de su planeador actual de manera que sus extremos cayeran un poco para hacerlas menos susceptibles a las ráfagas laterales –un rasgo más apropiado para los fuertes vientos de Kitty Hawk que para cualquier otro sitio–. Por primera vez, Orville empezó a pilotar el planeador. Se estrelló, la nave sufrió daños pero él no. Después de las reparaciones, volaron de nuevo, y superaron los planeos del año anterior. Sin embargo, el alabeo del ala no era todavía correcto, y a veces las colas fijas daban un problema nuevo. El planeador ya no hacía piruetas en torno a sus alas superiores, pero en cambio se deslizaba de lado en la dirección de las alas inferiores hasta que empezaba a dar vueltas en torno a ellas y se clavaba en la arena como un taladro. Los hermanos habían descubierto un nuevo peligro en el vuelo –la barrena de cola.

Las colas fijas eran demasiado eficaces. Las alas inferiores se movían ahora demasiado despacio, y por tanto tenían menos elevación que las alas superiores.

Su velocidad, por tanto, se reducía aún más, etc... hasta que la máquina giraba en la arena. Lo intentaron eliminando una de las colas fijas. Había poca diferencia. Después de una noche sin dormir, Orville dio con la solución: si el piloto pudiera controlar las colas verticales, entonces podría girarlas para aliviar la presión sobre sus caras inferiores. Podría equilibrar la elevación de las alas superiores e inferiores, y por tanto evitar que la nave entrara en barrena. Cuando estaban convirtiendo las paletas fijas en un timón controlable, tuvieron la idea posterior de conectar los cables que las controlaban a los que operaban el alabeo del ala. De esta forma, siempre que el piloto alabeara las alas, el timón giraría en la dirección apropiada.

La solución funcionó. Pudieron planear más de 500 pies [unos 160 metros] con un excelente dominio de la nave. Podían inclinarse y girar. Podían volar con el viento viniendo de lado. Y podían aterrizar con seguridad. Al fin habían establecido un sistema completo de control para un avión: el elevador trasero controlaba la subida y la bajada, y el alabeo de las alas y los timones verticales controlaban la inclinación y el giro. A finales de octubre, después de mucha práctica planeando, volvieron a Dayton. Antes de acabar 1902, habían solicitado una patente para el avión y su sistema de control.

* * *

Los Wright necesitaban un motor y unas hélices. Asumieron que los ingenieros de la marina habían resuelto el problema teórico de diseñar hélices. En realidad, no existía ninguna teoría. Por tanto, ¿cuál debería ser la teoría? Más tarde escribieron:

Lo que en principio parecía un problema sencillo se convirtió en más complejo cuanto más lo estudiábamos. Con la máquina moviéndose hacia adelante, el aire fluyendo hacia atrás, las hélices girando lateralmente y nada que se mantuviera fijo, parecía imposible encontrar un punto de partida a partir del cual examinar las diversas reacciones simultáneas. El análisis no estaba claro. Después de largas discusiones, nos encontramos a menudo en la absurda situación de que cada uno había cambiado a la posición del otro, sin mayor acuerdo que cuando

empezó la discusión.

Los hermanos desarrollaron su propia teoría. Una hélice mueve un barco al desplazar un volumen de agua. Corta el agua de la misma forma que un sacacorchos corta un corcho. Muchos de los rivales de los Wright supusieron que la hélice de un aeroplano debería ser una hoja plana que cortase el aire. Pero la analogía es falsa. A diferencia del agua, el aire es muy compresible. Por tanto, ¿qué forma debería tener una hélice? Los hermanos extrajeron una analogía brillante. Una hélice es un ala que se desplaza siguiendo una ruta en espiral. La analogía se basa en imaginar un modelo de un ala, llevar a cabo una rotación mental, y observar que corresponde a la mitad de una hélice giratoria. A medida que el ala gira, genera no elevación, sino empuje. Una hoja plana, como la de las hélices de Maxin o Langley, no genera mucho empuje. En vez de eso la hoja debe estar peraltada como un ala, y al empuje de las hélices se le aplican también los principios para la elevación.

Diseñaron una hélice para su máquina voladora. Era mucho más eficiente que cualquier otra hélice existente en cuanto a la conversión de energía en empuje. Y la teoría de los hermanos Wright predecía su empuje con un margen del uno por ciento. Estaban muy por delante de sus rivales, que continuaban apoyándose en el ensayo y error para los diseños de sus hélices. Después de haber construido y probado un prototipo, los hermanos construyeron dos hélices de madera. Usarían ambas, rotando en direcciones opuestas para equilibrar el par de torsión, y las montarían detrás de las alas de su máquina voladora para minimizar la turbulencia.

* * *

Los rivales de los Wright buscaron los motores más ligeros y potentes que pudieron conseguir. El motor de Langley, por ejemplo, era una maravilla de economía. Desarrollaba casi cincuenta y cinco caballos de potencia y pesaba poco más de 200 libras [unos 95 kilos]. Los Wright en una de sus maratonianas series de cálculos establecieron el mínimo con el que podrían tener suficiente. Estimaron el área de las alas de la máquina con motor (500 pies cuadrados) y su

peso total (625 libras)[2] incluidos el motor (200 libras) y el piloto (140-145 libras). Calcularon entonces la velocidad mínima que elevaría la máquina en el aire (veintitrés millas por hora) [treinta y siete km/h]. A partir de esta velocidad, calcularon la resistencia total de las alas y del área frontal de la máquina (noventa libras). Por tanto, el motor necesitaba producir noventa libras-fuerza de empuje. Y, finalmente, a partir de la velocidad y la resistencia, calcularon que el motor necesitaba tener ocho caballos de potencia. Escribieron a varias fábricas comerciales pidiendo un motor de combustión interna de 180 libras que pudiera generar de ocho a nueve caballos de potencia. Ninguna fábrica podía cumplir sus especificaciones con un coste razonable, y por tanto diseñaron ellos su propio motor.

Charlie Taylor, que era un habilidoso operario que trabajaba para ellos, lo construyó en su taller de bicicletas. El combustible se suministraba por gravedad; estaba vaporizado en un bidón de acero a través del cual pasaba el aire; la ignición se hacía mediante contactos de ruptura controlados por levas, y los cuatro cilindros se enfriaban con agua. Probaron el motor, pero al día siguiente la manivela se partió. Construyeron un motor nuevo. Pesaba menos de 180 libras, y liberaba casi doce caballos de potencia, más del mínimo.

* * *

Los hermanos tenían un motor y dos hélices, pero, ¿cómo podían conectarlos? El transmitir potencia de un motor a dos hélices que giraban en direcciones opuestas no era un asunto sencillo, escribió Orville. Los Wright eran expertos en la manipulación de modelos mentales de sistemas mecánicos, y puede que el lector sea capaz de adivinar la fuente de sus diseños para la transmisión. Una vez más, fue la bicicleta. Utilizaron una disposición de dientes y cadenas para transmitir la rotación del eje del motor a los dos ejes de las hélices. Para asegurarse de que una hélice girase en dirección opuesta a la otra, retorcieron la cadena una vez (de la misma forma que las correas de la Figura 2.5).

A mediados del verano de 1903, habían completado la transmisión, y estaban ocupados fabricando las partes de la máquina volante. Tenía la configuración canard, pero a mayor escala, con una envergadura un poco mayor de cuarenta

pies [unos doce metros]. La máquina se pasaba del peso estimado, pero calcularon que las alas serían suficientes para levantarla en el aire con la potencia del motor.

* * *

A finales de septiembre, partieron para su cuarto viaje a Kitty Hawk con su cuarta máquina, la máquina voladora con motor. Durante el mes siguiente pusieron a punto sus habilidades de vuelo, practicando con el planeador que se habían dejado el año anterior. Tuvieron noticias del vuelo inaugural de Langley. Su avión, el Great Aerodrome, había caído directamente en el río Potomac –un resultado que no les habría sorprendido si hubieran sabido que Langley había descuidado el control, y había ampliado a escala una maqueta de avión hasta alcanzar un tamaño suficiente para transportar a un hombre–. Su máquina voladora estaba lista, pero durante una prueba del motor se dañaron los ejes de las hélices. En nuevos ensayos del motor, uno de los ejes reparados desarrolló una pequeña fisura. Orville volvió a Dayton, y regresó con nuevos ejes hechos de acero macizo y las últimas noticias sobre el aeródromo de Langley. En su segundo vuelo, se rompió en el aire y cayó de nuevo en el Potomac. El piloto consiguió escapar de morir ahogado.

Con ayuda de los habitantes de la zona, los hermanos trasladaron la máquina voladora a su punto de despegue, donde se había colocado un raíl sobre el que correría. Wilbur ganó en el lanzamiento de la moneda, y por tanto fue el que hizo el primer intento. Arrancaron el motor, y, después de un problema inicial en la liberación del cable que la retenía, la máquina voladora corrió por el raíl. Antes de llegar al extremo de la sección final de este se movía más deprisa de lo que podía correr Orville. Wilbur empujó la nave hacia arriba, pero con demasiada inclinación. Alcanzó una altura de unos quince pies [unos cuatro metros y medio], pero luego se paró, depositándose con fuerza en el suelo y dañando los patines y algunas riostras. La repararon rápidamente. Pero el tiempo les retrasó un par de días. El jueves 17 de diciembre de 1903 despertaron con vientos de treinta millas por hora. Parecía ser su última oportunidad del año, y por tanto tendieron de nuevo la pista de despegue. Encontraron cuatro habitantes locales para ayudarles a trasladar la máquina voladora fuera del hangar. Arrancaron el

motor. Le tocaba a Orville volar el primero. A su señal, aproximadamente a las 10,35 a.m., la máquina voladora corrió por el raíl con Wilbur en la punta del ala derecha. La Figura 25.2 es la famosa fotografía de la máquina voladora nada más despegar. Se elevó en el aire y voló durante doce segundos bajo el control de Orville a una velocidad de unas treinta millas [cuarenta y ocho kilómetros] por hora. Recorrió una distancia de 120 pies sobre el suelo [treinta y seis metros y medio], y aterrizó en una posición que no era inferior a su punto de despegue. El mismo día se produjeron otros tres vuelos. En el más largo, Wilbur voló recorriendo 852 pies [260 metros] en 59 segundos.

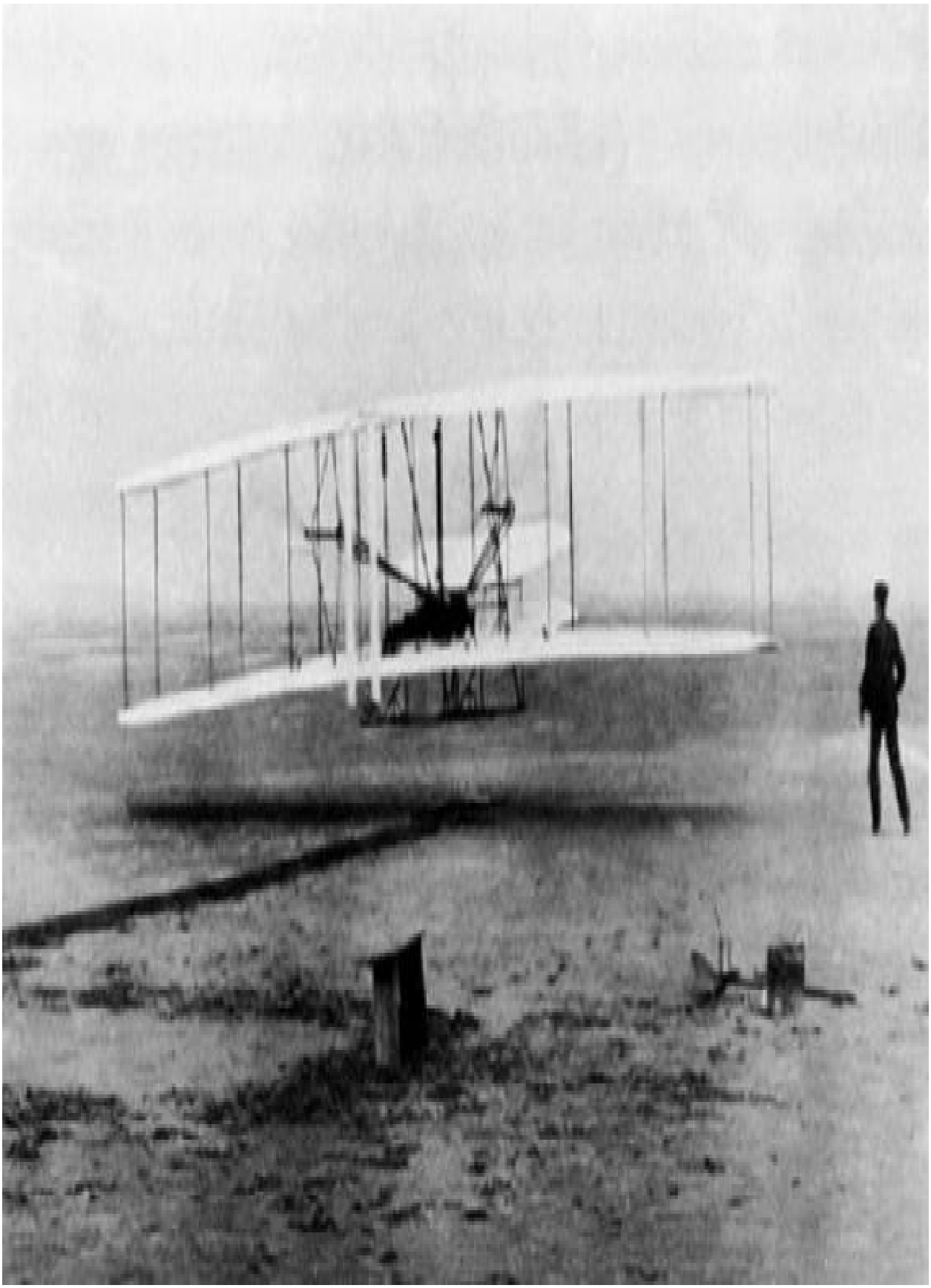


Fig. 25.2. La máquina voladora nada más despegar en su primer vuelo con éxito, el 17 de diciembre de 1903. Orville está tendido en el ala a los controles; Wilbur está a la derecha de la máquina. Son visibles las dos cadenas de bicicleta que transmiten la potencia desde el motor a las hélices.

En los años posteriores, los hermanos hicieron muchas mejoras en el diseño de sus máquinas voladoras, y separaron los controles para el timón y la inclinación de las alas. Hubo un parón de unos años durante los que no realizaron vuelos mientras trataban de vender sus máquinas, primero al Departamento de Guerra de Estados Unidos, y luego, cuando les rechazaron, a gobiernos europeos. Sus rivales no les alcanzaron. Wilbur llevó una máquina voladora a Francia y asombró a los aeronautas franceses, que se habían engañado a sí mismos considerándose a la cabeza de la carrera por volar. Orville hizo, finalmente, una demostración al ejército de Estados Unidos con una máquina voladora. Durante las pruebas, se rompió una hélice, y Orville se estrelló matando al pasajero que le acompañaba –la primera fatalidad en un aeroplano–. Su patente era correcta, no obstante se encontraron con largos litigios para hacer valer sus derechos. Ganaron todos los casos, pero Wilbur estaba agotado. Murió de tifus en 1912.

Orville vivió más tiempo. Fue testigo de casi todas las principales innovaciones en la aviación. El desarrollo de los alerones, aviones de una sola ala y motores potentes, que pronto dejaron atrás a su máquina voladora. Inventó el alerón [split], que todavía utilizan los aviones cuando aterrizan. Y comprobó que sus descubrimientos sobre el control de la nave y el diseño de hélices no se habían pasado de moda. No vivió lo suficiente como para ver un avance que le hubiera gustado. Fue el Gossamer Condor –un avión ultraligero diseñado por Paul MacCready–. Su hélice la movía no un motor, sino un ciclista muy entrenado que pedaleaba boca abajo en la aerodinámica cabina.

* * *

Poco después de iniciar su trabajo, los Wright se dieron cuenta de que el vuelo

no era un misterio que se pudiera resolver por ensayo y error. Exigía pensar de manera sistemática sobre muchos problemas. “No es asombroso”, escribió Orville, “que todos estos secretos hayan estado preservados durante tantos años solo para que nosotros pudiéramos descubrirlos”. ¿Por qué ellos consiguieron volar cuando tantos otros no lo lograron? Tuvieron algo de suerte, gran perseverancia, destreza con las manos, pero sobre todo, pensaron de manera excepcional. Sus rivales tenían prioridades erróneas; aumentaron a escala maquetas de juguete; trataron de volar sin entender los principios básicos. Los Wright volaron antes que sus rivales porque pensaron mejor que ellos.

Podemos dar sentido a su pensamiento, y podemos entender los mecanismos de los que depende. Eran expertos en encontrar contraejemplos para los argumentos: las consecuencias del fallo del motor mostraron que el control era más importante que la potencia. Sabían resolver inconsistencias creando posibles diagnósticos. Al resolver problemas, podían calcular, podían hacer deducciones y podían visualizar herramientas mecánicas que resolvían los problemas teóricos. Podían simular mentalmente el comportamiento de los procesos, como el del flujo de aire sobre el ala o la cola. Sobre todo, eran maestros de la analogía. Utilizaron cinco analogías fundamentales al inventar el aeroplano: la bicicleta mostraba que el control era más importante que la estabilidad; los buitres mostraban que el alabeo del ala haría que esta girara e inclinara la nave; una caja de cámara mostraba cómo alabear las alas sin perder la rigidez lateral; un ala giratoria mostraba cómo diseñar una hélice, y una cadena de bicicleta mostraba cómo transmitir la potencia del motor a las hélices. Eran genios de la visualización. Pero esta dependía de la habilidad para construir modelos mentales de entidades tridimensionales. Podían poner en movimiento estas representaciones para imaginar el flujo de aire sobre un avión, o para diseñar un sistema de transmisión. Y eran los más hábiles utilizando un modelo de una cosa, una bicicleta, como analogía de otra, un avión.

Notas al pie

[1] Nota de la tr.: El autor hace un juego de palabras entre have a wry humor [hacer gracia] y el nombre de los hermanos Wright.

[2] Nota de la tr.: 1 libra equivale a 453,5 g y 1 libra-fuerza equivale a 4,45 N.

CAPÍTULO 26

Desvelar un enigma

Durante la Segunda Guerra Mundial, un importante triunfo de los aliados fue descifrar el código Enigma de los alemanes. El código se basaba en una máquina, Enigma, que utilizaban todas las fuerzas alemanas. La capacidad de los aliados para leer los mensajes secretos entre el Alto Mando Nazi y las fuerzas aéreas, navales y terrestres alemanas tuvo un importante efecto en la aceleración del fin de la guerra en Europa. Varias personas y varios grupos de personas tomaron parte en esta historia –los criptoanalistas polacos, en especial Marian Rejewski, que fue el primero en descifrar el código, y Alan Turing, que había establecido la lógica abstracta de los ordenadores programables en 1936–. Formuló una teoría completa sobre Enigma para los descifradores de códigos británicos que trabajaban en la Escuela Gubernamental de Códigos y Cifrado [Governement Code and Cypher School] en Bletchley Park a unas cincuenta millas al norte de Londres.

Para los psicólogos, descifrar un código es un maravilloso ejemplo de cómo el razonamiento puede resolver un problema. La mayor parte de los que descifraron el código no nos dejaron ninguna explicación sobre cómo lo hicieron. Pero uno de los colegas de Turing, Gordon Welchman, que era también un brillante matemático, describió lo pasos de su pensamiento, aunque varios años después. Descifró el código durante los meses de octubre y noviembre de 1939, poco después de que la guerra hubiera comenzado. Trabajaba solo, y sin saber nada de sus precursores. Cuando informó de su éxito, su jefe en Bletchley Park se puso furioso porque el código ya había sido descifrado. Welchman se enteró de que otros –los criptoanalistas polacos– lo habían descifrado, y de que ya estaba en marcha un sistema de tarjetas perforadas similar al que él proponía. Sin embargo, su trabajo para resolver el Enigma le ayudó a prepararse para organizar el tráfico masivo de decodificaciones. Y su explicación nos permite comprender cómo es la mente extraordinaria de alguien que descifra códigos de forma satisfactoria.

El razonamiento de Welchman se basaba en su conocimiento de la máquina Enigma y del protocolo alemán para su uso. La parte central de su pensamiento fue una secuencia ejemplar de deducciones, incluidas algunas que se basaban en probabilidades, y por ello voy a describir cómo aclara la teoría de modelos de la deducción sus procesos mentales. Para entender estos procesos, sin embargo, es necesario conocer en primer lugar lo que Welchman ya sabía cuando empezó a pensar en el problema. En consecuencia, empiezo con una descripción de los códigos y de la máquina Enigma, y del protocolo alemán para configurar la máquina antes de codificar o decodificar un mensaje. Fue en ese punto donde cometieron un error fatal. Y posteriormente analizo el pensamiento de Welchman mientras descifraba el código.

* * *

Al inicio de la Segunda Guerra Mundial, la estrategia alemana de Blitzkrieg –guerra relámpago– exigía que los generales lucharan cerca de la primera línea de sus tropas móviles. El éxito de esta estrategia dependía de un método de comunicación rápido y seguro. Los mensajes no solo eran un secreto para el mundo exterior, sino también para cada unidad respecto a las demás fuerzas terrestres, aéreas y navales, el ejército privado de Hitler (las SS) y las unidades de la inteligencia alemana. Los alemanes utilizaban máquinas Enigma para encriptar sus mensajes secretos. Los panzer del general Guderian, por ejemplo, estaban equipados con estas máquinas durante la invasión de Francia en 1940. La Figura 26.1 muestra una fotografía del general en un vehículo armado de transporte durante esa invasión, y en primer plano a la izquierda de la foto se puede ver una máquina Enigma.

Los códigos son casi tan viejos como la guerra. César Augusto se basó en un código “César” sencillo en el que para codificar un mensaje, las letras del alfabeto se desplazaban un número fijo de letras. Si el desplazamiento es, pongamos por caso, tres letras, entonces la A se codifica como D, la B como E, etc. Debido a que el código es fácil de resolver, un procedimiento mejor es asociar cada letra a una letra elegida al azar, p. ej., la A se codifica como W, la B como A, etc., hasta que a cada letra se le asigna un código diferente. Si ha leído usted la historia corta de Edgar Allan Poe, El escarabajo de oro, entonces sabrá

que este tipo de código también es fácil de resolver. Hacemos corresponder la frecuencia con que aparece cada letra en un conjunto de mensajes encriptados con las frecuencias conocidas de las letras en el idioma original, o “texto claro”, de los mensajes. La letra más frecuente en los textos en inglés es la E. Por consiguiente, si la letra más frecuente en los mensajes encriptados a partir del inglés es la S, podemos inferir que en el cifrado la E se encripta como S. La segunda letra más frecuente se puede decodificar en inglés de manera similar como T, etc. Solo necesitamos conocer las frecuencias relativas con que aparecen las letras del alfabeto en el lenguaje escrito de los textos claros.

Una forma de obviar un análisis de frecuencia es utilizar un “cifrado de uso único”. El cifrado utiliza un código César diferente para cada letra de un mensaje. La primera letra del texto claro, sea cual sea, se codifica, pongamos por caso, como la que está cinco letras después, la segunda letra como la que está 22 letras después, etc., hasta completar el mensaje. El cifrado de uso único, que determina la codificación, tiene que ser por tanto tan largo como el texto claro del mensaje, y cada aparición de una letra concreta, como la E, suele tener una codificación diferente. El cifrado es indescifrable excepto por la fuerza bruta, es decir, intentando cada posible codificación para cada posible letra. Si ninguna letra se codifica como ella misma, entonces en un mensaje hay veinticinco posibles codificaciones para cada letra. Dado un mensaje de, digamos, veinte letras, quienes descifren el código tendrán que examinar en el peor de los casos un gran número de posibles codificaciones (2520) para resolver el mensaje. Incluso si pudieran examinar cada posible codificación en una millonésima de segundo, les podría llevar casi trescientos millones de millones de años descifrar el código. El código es indescifrable pero con un coste. Tanto quien codifica el mensaje como quien lo decodifica deben tener el mismo cifrado de uso único, y mucho tiempo, tanto para cifrar como para descifrar mensajes. El general Guderian no podría haberse comunicado con sus tropas de esta forma. Lo que necesitaba era un sistema de encriptado rápido pero impenetrable. Lo que necesitaba era una máquina que hiciera el trabajo.

Fig. 26.1. El general Guderian, de pie con prismáticos de campo, en un vehículo armado de transporte en 1940 durante la invasión de Francia. En primer plano están dos oficiales de cifrado y un radiooperador. La máquina Enigma (con tres rotores) aparece en primer plano a la izquierda.

La Figura 26.2 es una foto de una máquina Enigma. Parece una combinación de una máquina de escribir portátil y una antigua calculadora de mesa. En su parte frontal pueden verse cables conectados en diversos enchufes. El teclado tiene teclas para las veintiséis letras mayúsculas del alfabeto. Tras ellas hay un conjunto de luces, una para cada letra del alfabeto con la misma distribución espacial de las teclas del teclado. Hacia la parte trasera de la máquina hay tres rotores (más tarde, durante la guerra, las máquinas tenían cuatro rotores), La característica central de la máquina es que no se apoya en un sistema constante de codificaciones. Al escribir una letra, como por ejemplo una Q, lo que se enciende es alguna otra letra, como por ejemplo una V. Al escribir de nuevo una Q, se enciende otra letra diferente, como por ejemplo una E. La siguiente vez que se escriba una Q se enciende de nuevo una letra diferente, etc. Enigma está protegida por tanto frente al criptoanálisis basado en las frecuencias de aparición de las diferentes letras. El método es inútil, porque en Enigma la correspondencia entre las letras del texto claro y sus versiones encriptadas cambia cada vez que se escribe una letra en el teclado. La máquina es un cifrado de uso único automatizado.

Fig. 26.2. Una máquina Enigma y sus dos rotores extra (© Deustches Museum, 80306 Munich, Alemania).

Las codificaciones son simétricas, es decir, si al escribir una Q se enciende la letra V, entonces si se hubiera escrito una V, se habría encendido la letra Q. Esta propiedad era crucial para la decodificación de los mensajes. El encargado de la codificación configuraba su máquina con una posición determinada y codificaba un mensaje. La primera letra del texto claro del mensaje original era una Q, y por tanto escribía la Q y se encendía la letra V. Esta era la primera letra del mensaje codificado. El encargado de la decodificación configuraba su máquina con la misma posición inicial y escribía el mensaje encriptado en su máquina. Lo que se encendía era la secuencia de letras en el texto claro original. La primera letra del mensaje codificado era una V. Escribía V y se encendía la letra Q. Esta era la primera letra del texto claro del original. La simetría de la máquina aseguraba que la codificación y la decodificación fueran operaciones simétricas.

La máquina dependía de dos componentes independientes. El teclado de conexiones de la parte delantera de la máquina permitía al encargado del cifrado hacer conexiones arbitrarias entre las teclas y las letras que se encendían. Debido a la necesidad de simetría, cuando el operador utilizaba los cables especiales para conectar el teclado de conexiones, se creaba una relación simétrica, es decir, si la tecla A encendía la luz R, entonces la tecla R encendía la luz A. Había once pares de letras simétricos. Las restantes cuatro letras del alfabeto no estaban afectadas por la conexión del teclado de conexiones. En realidad, estaban conectadas a sí mismas. Había más de 200 millones de millones de formas en las que el operador podía conectar el teclado de conexiones. Una vez que lo había conectado, permanecía sin cambios durante veinticuatro horas.

El segundo componente era el sistema de rotores en la parte trasera de la máquina, conocido como scrambler. Tenía terminales para las veintiséis letras. Cuando el operador pulsaba la tecla de una letra, la corriente circulaba desde la batería a través de la tecla, y luego a través del teclado de conexiones hasta el terminal adecuado del scrambler. Atravesaba el scrambler por una ruta compleja que describiré dentro de un momento, cambiaba de dirección y salía por un camino diferente hasta que alcanzaba un terminal diferente del scrambler.

Atravesaba de nuevo el teclado de conexiones (por un camino diferente y en dirección opuesta a la que tenía al entrar) y encendía una de las luces de las letras. Esta letra era la codificación de la letra que el operador había tecleado. La Figura 26.3 muestra esta ruta.

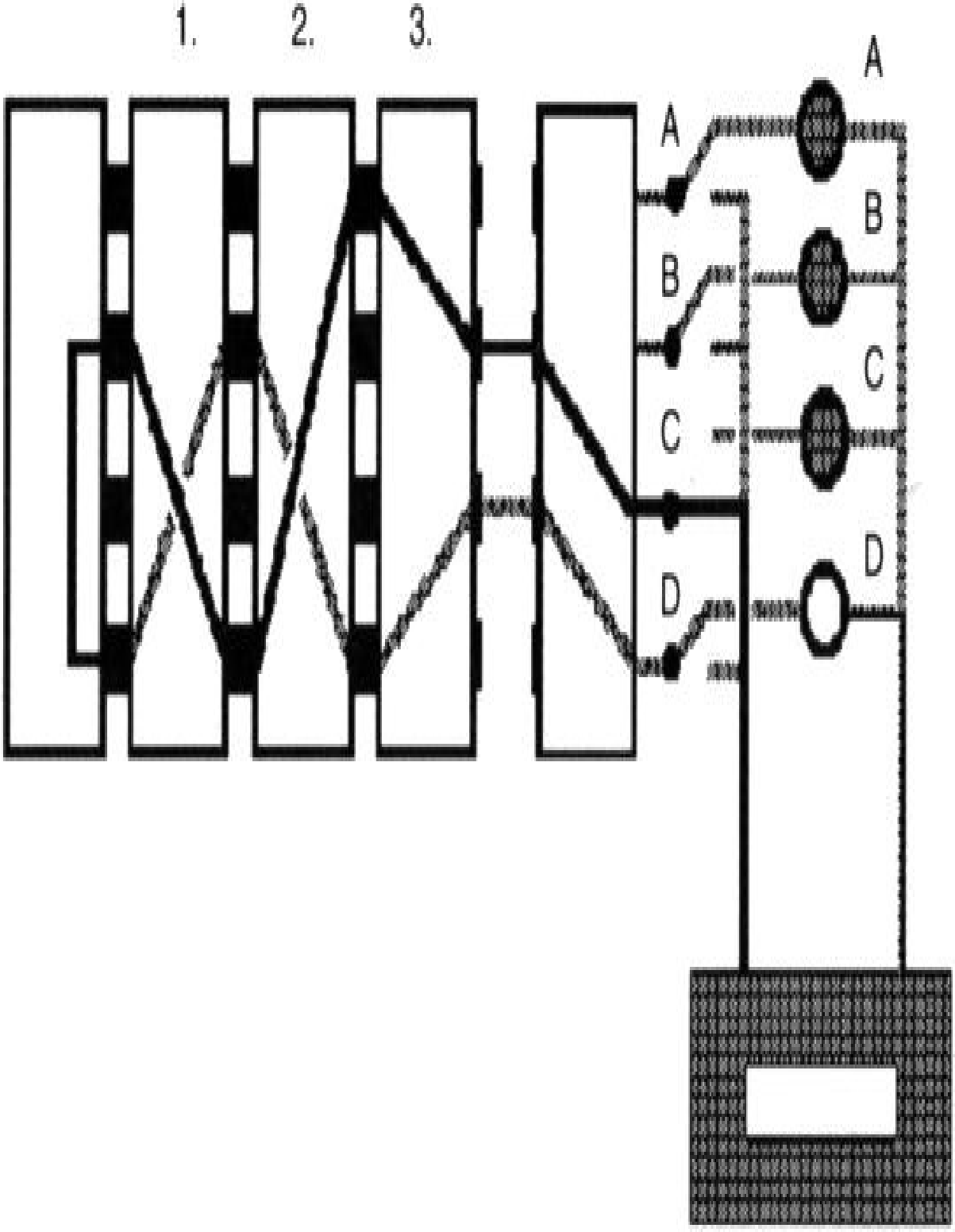

1.
2.
3.
A
B
C
D
A
B
C
D

Fig. 26.3. Un diagrama de las conexiones de una máquina Enigma. Muestra solo cuatro teclas y cuatro luces. Se ha pulsado la tecla C, y el diagrama muestra el flujo de corriente que completa el circuito para encender la letra D.

Los terminales en la entrada del scrambler estaban dispuestos en círculo. Estaban en contacto físico con el rotor contiguo. Cada rotor estaba conectado de manera que la entrada de una letra llevaba a una salida que correspondía a una letra diferente. La Figura 26.4 es un diagrama simplificado de un único rotor, mostrando solamente las conexiones para cuatro letras del alfabeto. El rotor que recibía la entrada era adyacente al siguiente rotor, que estaba conectado de manera análoga, y así una y otra vez, hasta que la corriente regresaba mediante un reflector estático y pasaba a través de los rotores en orden inverso hasta que aparecía para encender una letra.

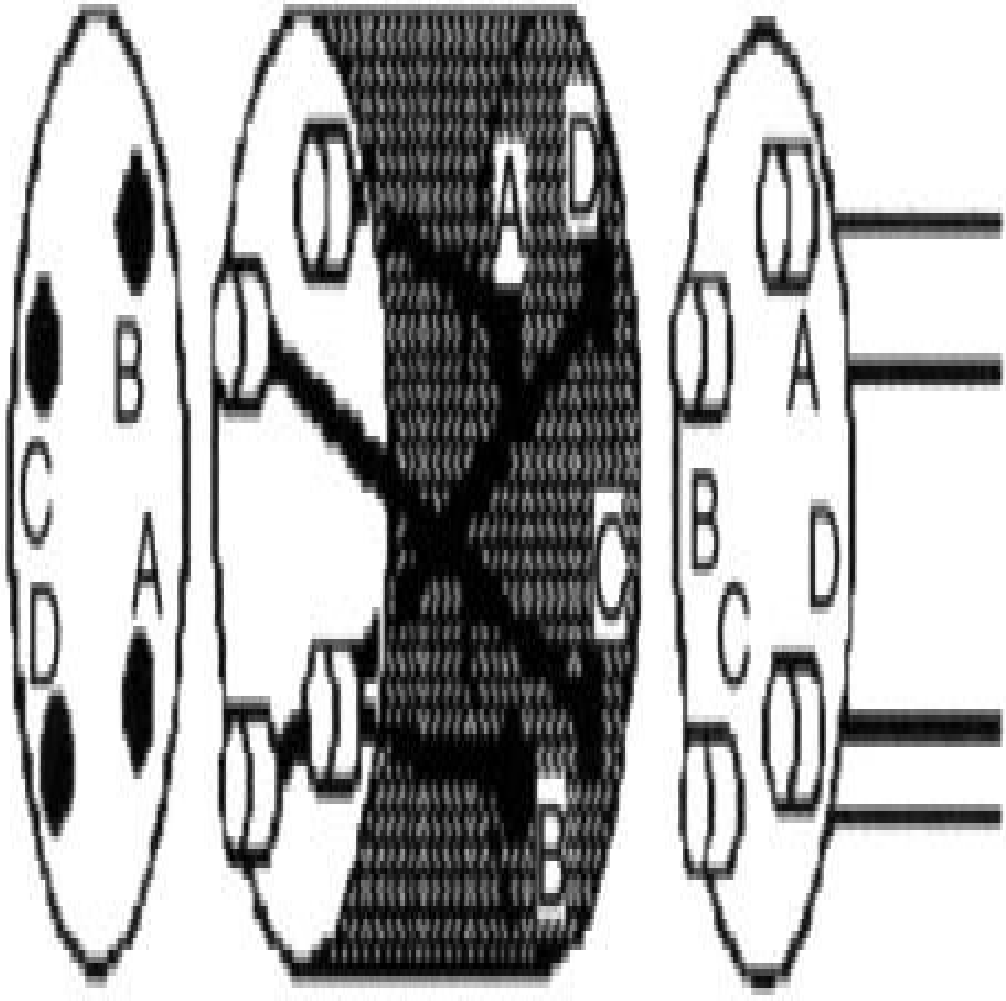

B
C
A
D
A
D
C
B
A
B
D
C

Fig. 26.4. Un diagrama simplificado de las conexiones de un único rotor, mostrando solo cuatro letras. Cada terminal de entrada en el rotor está conectada a un terminal de salida diferente. El teclear como entrada C, por ejemplo, corresponde a la salida de B en el rotor y desde allí a A en el siguiente rotor.

El scrambler era por tanto una herramienta que recogía como entrada una letra y daba otra letra como salida. Cada uno de los tres rotores tenía unas conexiones cruzadas diferentes entre la posición de una letra y la de otra. En consecuencia, había una compleja serie de correspondencias a partir de la entrada del primer rotor, de este rotor al segundo, etc., a lo largo de todo el recorrido hasta la salida. Cuando el operador pulsaba una tecla, el primer rotor giraba una posición, con lo que ahora había una nueva serie de rutas para atravesar el scrambler. El primer rotor seguía girando un paso con cada nueva tecla pulsada hasta que alcanzaba una posición, parecida a la de un cuentakilómetros, en que tanto él como el segundo rotor giraban un paso. De manera similar, el segundo rotor finalmente alcanzaba una posición en la que arrastraba al tercer rotor, y los tres rotores giraban a la vez. En la máquina en uso al inicio de la guerra, los rotores no regresaban a sus posiciones iniciales hasta que habían recorrido un gran número de posiciones, correspondientes al mismo número de pulsaciones (26 x 26 x 26 = 263 = 17.576). Cada posición daba lugar a un encriptado diferente del alfabeto.

La posición inicial del rotor al comienzo de la codificación de un mensaje no tenía que corresponder a la letra A. Unido al rotor había un anillo independiente que contenía las letras en orden alfabético, y el encargado del cifrado podía girar este anillo y situarlo en posición con el fin de asignar la posición inicial del rotor a cualquier letra del alfabeto. Todos los rotores tenían estos anillos.

En resumen, la máquina Enigma tenía dos componentes independientes que creaban sus codificaciones. Un teclado de conexiones que se enchufaba a mano antes de codificar o decodificar mensajes, y un conjunto de tres rotores –más tarde, durante la guerra, cuatro rotores–, que producían también un código aleatorio con la propiedad de que cada vez que se pulsaba una tecla, el scrambler rotaba para cambiar el código de la siguiente letra. Con el fin de producir un número aún mayor de posibles codificaciones, los tres rotores del scrambler se

escogían de un conjunto de cinco, y podían insertarse en la máquina en cualquier orden. Estas posibilidades creaban una cantidad enorme de situaciones diferentes (sesenta posibles conjuntos de tres rotores multiplicadas por 263 posiciones para los anillos). Estas posibilidades multiplicadas por las de las conexiones del teclado de conexiones daban un número total de posibles códigos de más de doscientos millones de millones de millones. Este número era lo suficientemente grande como para hacer que Enigma pareciera inexpugnable.

* * *

Para enviar un mensaje desde el alto mando en Berlín al general Guderian sentado en su vehículo blindado había que seguir una serie de pasos que aseguraran el éxito. La máquina de Berlín y la máquina del vehículo tenían que estar configuradas en idénticas posiciones iniciales. Sus teclados de conexiones tenían que estar conectados de la misma forma; tenían que tener en el scrambler los mismos rotores situados en las mismas posiciones; los anillos de letras de esos rotores tenían que estar fijados sobre ellos en las mismas posiciones, y los rotores con sus anillos en posición tenían que girarse hasta las mismas posiciones iniciales en ambas máquinas. Solo si esta configuración se llevaba a cabo correctamente, ambas máquinas tenían el mismo estado inicial. El operador en Berlín escribía el texto claro en su máquina, y la resultante secuencia de letras iluminadas se transmitía por radio en código morse. El operador de radio en el vehículo blindado de Guderian transcribía la secuencia, y la pasaba al encargado del cifrado que estaba sentado a su lado. El encargado del cifrado tecleaba la codificación en su máquina, y la secuencia resultante de las letras iluminadas, si no se habían producido errores, era el texto claro del mensaje original de Berlín. En la Figura 26.1 se pueden ver en primer plano al radiooperador y a dos encargados del cifrado.

¿Cómo podían los encargados de la emisión y recepción del cifrado configurar sus máquinas independientes de manera que tuvieran un estado inicial idéntico? La respuesta dependía tanto de un libro de códigos que ambos operadores poseían como de las señales iniciales de la radio transmisión. Los libros de códigos, que variaban de una unidad a otra, contenían una clave que especificaba qué tres rotores utilizar, su orden en el scrambler, las disposiciones de sus tres

anillos con el alfabeto y las conexiones del teclado de conexiones. La clave era válida durante veinticuatro horas, y por tanto sería utilizada para muchos mensajes, aunque estaba limitada a un determinado tipo de unidad militar. Welchman había estudiado el tráfico de radio, y sabía que los alemanes estaban trasmitiendo al menos cien mensajes al día utilizando una determinada clave. Cada clave tenía un número, y este número se transmitía en texto claro al comienzo de cada mensaje de radio, de manera que el encargado de la recepción podía comprobar si estaba en su libro de códigos. Si no estaba, el mensaje no tenía importancia para su unidad, y no podría decodificarlo. Por descontado, los analistas británicos podían reunir todos los mensajes enviados utilizando una determinada clave.

Lo único que la clave no especificaba eran las posiciones iniciales de los rotores y sus anillos al comienzo de un mensaje. Este tema se dejaba a discreción de los encargados de la codificación, quienes se suponía que escogían una nueva posición al azar para cada mensaje. (La pereza de los operadores iba a convertirse en una gran ayuda para quienes descifraban los códigos.) El operador de Berlín decidía las posiciones iniciales de los rotores que utilizaría en la codificación de un mensaje para Guderian. Supongamos que decidía que las posiciones iniciales de los tres rotores debían ser, respectivamente: MCR. Tenía que girar a mano los rotores de forma que pudiera ver la letra inicial M en el primer rotor, la C en el segundo rotor y la R en el tercer rotor. Estas letras eran visibles en unas ventanitas próximas a cada rotor (véase la tapa abierta con tres ventanas en la Figura 26.2). Luego tenía que transmitir estas posiciones de los tres rotores de manera que el encargado de la recepción, que ya tendría su máquina en la configuración adecuada, pudiera girar los rotores de su máquina a las posiciones iniciales correctas. Pero, si el operador transmitía sus posiciones iniciales en texto claro, podría descubrir información vital. Sin embargo, no podía transmitirlas en un texto cifrado, porque formaban parte de la configuración inicial necesaria para el cifrado. Por tanto, ¿cómo se hacía?

El operador de Berlín había elegido una disposición inicial de los rotores para codificar un mensaje que era MCR. Para comunicar esta disposición, giraba los rotores hasta una posición al azar, digamos, NIV. Esta posición se comunicaba por radio en texto claro antes del texto cifrado. A partir de esta posición de NIV, tecleaba luego dos veces la disposición inicial de los rotores para el cifrado: MCRMCR. El resultado era, pongamos por caso, esta codificación:

PUESQW

En la transmisión de radio aparecía, a continuación, esta secuencia. El encargado de la transmisión colocaba entonces sus rotores en la posición MCR, y encriptaba el mensaje, que se añadía a la señal de radio.

El operador de Guderian recibía, tras los signos de llamada y la identidad de la clave, las siguientes secuencias de letras:

NIV – el texto claro de la posición elegida al azar por el encargado del envío para transmitir la disposición inicial de los rotores para decodificar el mensaje principal.

PUESQW – el encriptado de la disposición inicial de los tres rotores repetida dos veces.

El texto cifrado del mensaje a partir de la disposición inicial.

El operador de Guderian ya había configurado su máquina de acuerdo con la clave del día. Giraba sus rotores a la posición NIV y luego tecleaba la cadena PUESQW. Esta cadena iluminaba las letras MCRMCR, y por tanto ahora el operador conocía la disposición de los tres rotores necesaria para decodificar el mensaje encriptado. Giraba sus rotores a la disposición MCR, y tecleaba el mensaje encriptado, que la máquina decodificaba. Solo un operador que hubiera configurado su máquina de manera adecuada podría decodificar PUESQW para revelar las disposiciones iniciales de los rotores para la decodificación del mensaje principal.

Ya sabe usted tanto como sabía Welchman. Es decir, usted sabe todo lo que es esencial para entender cómo descifró el código Enigma, o incluso para tratar de

adelantarse a la solución que él dio.

* * *

Supongamos que una máquina Enigma no tuviera rotores en un scrambler, sino solo el teclado de conexiones. Podría producir millones y millones de códigos diferentes, pero descifrar cualquiera de ellos sería un juego de niños utilizando un análisis según el cual la letra más frecuente del mensaje correspondería a la letra más frecuente en el idioma del texto claro, etc., como describí anteriormente. Supongamos además que no hubiera más que 250 posibles configuraciones para una máquina Enigma. Sería factible ensayarlas una tras otra para descubrir cuál de ellas conduce a una decodificación satisfactoria del mensaje, es decir, a un mensaje sensato en alemán obtenido mediante un análisis de frecuencias. Y por tanto lo central en el problema es la tarea de reducir el número de posibles configuraciones de la máquina a un número manejable. Este procedimiento es ciertamente el que siguió Welchman. Describió sus reflexiones en una secuencia de pasos, que yo voy a seguir.

Paso 1

¿Qué pista hay en la forma en que los alemanes utilizaban la máquina que pudiera ayudar a reducir el gran número de posibles códigos? ¿Están descubriendo alguna información sobre la configuración inicial de la máquina? Algunas veces, cuando jugamos a un juego como por ejemplo el ajedrez, tenemos una poderosa intuición respecto al movimiento correcto que debemos realizar incluso si en principio no somos capaces de ver por qué es el movimiento correcto. (Describí estas intuiciones en la Parte II del libro.) A menudo presagian una solución. El primer paso de Welchman fue una intuición de este tipo. Sospechaba, como bien puede haber hecho usted, que la clave podría estar en la doble codificación de la disposición inicial de los rotores del scrambler. La redundancia en el protocolo de los alemanes estaba diseñada para minimizar errores. Dado que el operador de Berlín había escogido MCR como

disposición inicial, estaba obligado luego a transmitir su doble codificación a partir de alguna otra posición arbitraria de los rotores. El encriptado resultante, PUESQW, significa que la configuración de la máquina codificaba inicialmente una letra desconocida como P y luego tres pulsaciones más tarde como S. Existen relaciones similares entre los otros pares de letras: U y Q, y E y W. Estas inferencias son sencillas deducciones a partir de un modelo mental de la máquina.

Paso 2

El segundo paso en el pensamiento de Welchman se basó en un examen de estas dobles codificaciones para las disposiciones iniciales. He aquí algunos ejemplos típicos. Debería usted leerlos en voz alta, y ver si nota alguna diferencia entre el primero y los dos últimos:

PUESQW

SPESNT

VBYQGY

Espero que se haya dado cuenta de que cada una de las dos últimas cadenas contiene una letra repetida. La propiedad crítica, sin embargo, no es solo el que aparezca dos veces la misma letra, sino sus posiciones. En la cadena S PE S NT, la letra S aparece en las posiciones primera y cuarta. Por consiguiente, la máquina repite una letra para codificar la misma letra desconocida al describir la disposición inicial de los rotores. Esta letra repetida aparece con una distancia de tres pulsaciones. De la misma forma, la cadena VB Y QG Y codifica la tercera

letra al describir la disposición inicial desconocida utilizando en ambos casos la Y, con una distancia de tres pulsaciones. Me referiré a este tipo de codificaciones, es decir, a aquellas para las que la máquina repite la misma letra tres pulsaciones después cuando codifica dos veces la misma letra, como repeticiones. La detección de repeticiones por parte de Welchman fue inductiva: examinó muchas dobles codificaciones, y se dio cuenta de que algunas de ellas contenían repeticiones. Cuenta que empezó a pensar, por casualidad, sobre estas repeticiones. El segundo paso culminó en una sutil cuestión: ¿era siempre posible generar una repetición a partir de cualquier posición que se eligiera para los rotores del scrambler? Es decir, ¿cualquier configuración de la máquina da siempre la misma codificación para al menos una letra del alfabeto después de codificar en medio otras dos letras?

Paso 3

Si no siempre es posible una repetición, entonces la aparición de una repetición excluye un cierto número de posibles configuraciones de la máquina. Esta conclusión exige una inferencia más sutil:

Supongamos que hay una propiedad de un encriptado, que llamaré propiedad X, que es una consecuencia de la configuración de la máquina. Si todas las configuraciones permiten que aparezca X en algunos mensajes, entonces la aparición de X en un mensaje es compatible con cualquier colocación. Pero, si algunas configuraciones no permiten que se produzca X en ningún mensaje, entonces la aparición de X en un mensaje excluye esas configuraciones, es decir, X nos permite eliminar esas configuraciones. Una repetición es una propiedad del tipo de X. Por consiguiente, si algunas configuraciones no permiten que aparezcan repeticiones en ningún mensaje, entonces la aparición de una repetición en un mensaje excluye esas configuraciones.

La deducción que está en el centro de este razonamiento gira sobre premisas que contienen varios cuantificadores, los cuales he marcado en cursiva. Hacen difícil

la deducción, pero un razonador habilidoso puede utilizar modelos para hacer estas inferencias (véase Capítulo 11).

Welchman descubrió que las repeticiones no podían aparecer en todas las posibles disposiciones. Examinó ejemplos de codificaciones para posiciones sucesivas de una máquina, y vio que en algunas configuraciones no aparecían repeticiones con una distancia de tres pulsaciones para ninguna letra del alfabeto. Este fue su tercer paso: las repeticiones no pueden producirse siempre, y por tanto la aparición de una repetición excluye posibles configuraciones. Como él escribió, este hecho fue el germen a partir del que iba a desarrollarse su éxito.

Paso 4

Su siguiente paso fue encontrar la probabilidad de que se produjera una repetición en la doble encriptación de la disposición inicial de los rotores. Welchman nos cuenta que el cálculo de probabilidades no era uno de sus puntos fuertes, pero que los cálculos eran elementales. Podían llevarse a cabo a partir de modelos mentales de probabilidades y de asumir la equiprobabilidad (véase el Capítulo 15). En cada posición de los rotores en el scrambler, la máquina Enigma produce una codificación simétrica basada en trece pares de letras. Consideremos dos posiciones diferentes de los rotores, y supongamos que en la primera de ellas la A se codifica como R (y por tanto la R se codifica como A). El conjunto completo de parejas en la primera posición es, pongamos por caso:

A B C D E F H I J K M N V

| | | | | | | | | | | | |

R G L Y P Z U W T S O Q X

En la segunda posición de los rotores, la probabilidad de que A se codifique de nuevo como R es una entre veinticinco, porque hay veinticinco letras diferentes

de la A entre las que elegir su codificación. El mismo argumento se aplica a cada uno de los trece pares, es decir, cada uno de ellos tiene una probabilidad de repetición de 1/25. Así que, las posibilidades de que al menos un par se repita en la segunda posición de los rotores son por tanto de trece sobre veinticinco.

La selección de cualquier par para codificar la disposición inicial del primer rotor es una entre trece. Por consiguiente, la probabilidad de que al codificar la disposición del primer rotor se produzca una repetición es igual a la probabilidad de que al menos un par se repita (13/25) multiplicada por la probabilidad de seleccionar ese par para codificar la posición de la posición del primer rotor (1/13), es decir, la probabilidad es 1/25. Las posibilidades de una repetición en las posiciones dos y cinco son las mismas, como lo son las posibilidades de una repetición en las posiciones tres y seis. En resumen, las posibilidades de que aparezca una repetición son tres entre veinticinco, o aproximadamente una entre ocho. El cuarto paso, por tanto, mostró que las posibilidades de una repetición al codificar la disposición inicial de los rotores deberían ser aproximadamente una entre ocho.

Paso 5

Este paso fueron una serie de deducciones. Empezaba con la inferencia de que la probabilidad de una repetición depende de las posiciones de los rotores (y sus conexiones internas cruzadas). No se ve afectada en forma alguna por la elección de las conexiones con el teclado de conexiones. Esta elección cambiará la letra concreta que se repite, pero no las posibilidades de una repetición. Se sigue a continuación que los 200 millones de millones de posibles formas de conectar el teclado de conexiones no tienen efecto sobre la aparición de repeticiones. Lo único que importa son los rotores. Como expliqué anteriormente, hay sesenta opciones para escoger los tres rotores entre un conjunto de cinco y hay 17.576 disposiciones distintas para los anillos con el alfabeto. Las deducciones de Welchman, por tanto, reducían el número de posibles configuraciones a analizar al producto de estos dos números, es decir, poco más de un millón.

Un preámbulo típico de un mensaje Enigma conteniendo una repetición era:

KCX – las posiciones iniciales arbitrarias en el texto claro.

SPENST – la doble codificación de la consiguiente disposición de los rotores para la codificación o decodificación del mensaje.

Los analistas saben que los rotores fueron girados a la posición KCX, pero no conocen la configuración del scrambler –la elección de los rotores o las disposiciones de sus anillos con el alfabeto–. En teoría, podrían teclear en una máquina KCX para cada una de las aproximadamente un millón de situaciones. Alrededor de la mitad de las posiciones darían la repetición requerida de una letra en las posiciones primera y cuarta. Aquellas posiciones que no producen una repetición se pueden eliminar. Se utiliza la misma clave durante veinticuatro horas, y cada mensaje contiene el texto claro de la posición de partida arbitraria para los tres anillos. Se sigue que cada repetición en el preámbulo de un mensaje nos permite eliminar alrededor de la mitad de las posibilidades partiendo del conjunto de poco más de un millón de posibles disposiciones de los rotores y los anillos. La aritmética elemental dio lugar a la inferencia siguiente: después de la primera repetición, solo quedan poco más de medio millón de posibles configuraciones. Después de la segunda repetición, hay solo poco más de un cuarto de millón. Y después de doce repeticiones, quedan solo 250 posibles configuraciones. Se podría examinar cada una de ellas para comprobar si era la disposición correcta, pero esta tarea era factible.

Paso 6

Una repetición, como mostró el paso 4, debería producirse por término medio en uno de cada ocho mensajes. En cien mensajes, debería haber una media de doce repeticiones, y los alemanes estaban transmitiendo cada día al menos cien mensajes para una determinada clave. Por tanto, el sexto paso concluía con la inducción de que debería haber mensajes suficientes para reducir el número de posibles configuraciones de la máquina que está utilizando una determinada clave a unas 250.

Hemos seguido ya los seis pasos del pensamiento de Welchman que le llevaron a averiguar cómo descifrar el código. Debería usted haber comprendido que cada aparición de una repetición para una determinada clave permite que el número de posibles disposiciones iniciales se reduzca a la mitad, y que deberían producirse repeticiones suficientes como para permitir reducir las posibilidades a un número manejable. Las restantes ideas de Welchman se referían a la implementación práctica de su solución en una era previa a los ordenadores. Su idea era examinar de una vez por todas si se producía una repetición en cada una de las aproximadamente un millón de posibles disposiciones de una máquina y almacenar los resultados en grandes hojas perforadas. Los agujeros de las hojas representaban las repeticiones. Las hojas podían apilarse sobre una mesa iluminada y alinearse de una manera adecuada. Cada vez que una luz atravesara todas las hojas, indicaría una posible disposición del anillo con el alfabeto y de la distribución de rotores que era utilizada en la codificación. La comprobación de una máquina y el perforado de las hojas llevaría varios meses. Pero, una vez hecho, se podría utilizar todos los días. El apilamiento de las hojas y el aislamiento posterior de la posición correcta entre las 250 podría llevar medio día. Welchman calculó que el trabajo se haría con una media de 760 apilamientos. Una vez conocida la configuración de una máquina, podría utilizarse para teclear el mensaje codificado. El resultado era un cifrado estático invariable basado solamente en el teclado de conexiones, que podría descifrarse utilizando un análisis de frecuencia del tipo del que describí anteriormente.

* * *

Cada paso en el razonamiento de Welchman se puede explicar en función de los principales procesos de razonamiento inductivo y deductivo que se explican mediante la teoría de modelos. Sus principales inducciones se basan en observaciones del resultado de salida de las máquinas. Todas sus deducciones principales podrían simularse en un programa de ordenador equipado con un modelo en funcionamiento de una máquina Enigma. La Tabla 26.1 resume sus seis pasos desde este punto de partida. Cada paso puede explicarse en función de los procesos básicos de inferencia. Lo que presenta un reto mucho mayor es diseñar un programa de ordenador que descubriera por sí mismo la secuencia de pasos. Welchman necesitó una semana de pensamiento intenso y continuado para

resolver el problema. Cada paso preparaba el camino para el siguiente. Trabajaba avanzando hacia la meta de la misma forma en que nosotros nos enfrentamos a problemas simples de ingeniería inversa. Pero sospecho que un análisis medios-fines dio lugar a la fragmentación del problema en dos componentes: la reducción de las disposiciones del scrambler a un número factible y la utilización de un análisis de frecuencia para resolver las conexiones del teclado de conexiones.

Tabla 26.1. Los principales pasos en el razonamiento de Welchman

1. La doble codificación de las posiciones iniciales del rotor podría ser útil: una intuición inconsciente.

2. Al menos en algunas configuraciones se producen repeticiones: una inducción. Si en otras configuraciones las repeticiones son imposibles, entonces la aparición de una repetición excluye esas configuraciones: una deducción.

3. Las repeticiones son imposibles en algunas configuraciones del scrambler: una inducción.

4. La probabilidad de una repetición es aproximadamente de una por cada ocho mensajes: una serie de deducciones.

5. Esta probabilidad depende solamente de la posición del scrambler. Por tanto, el análisis puede prescindir de las conexiones del teclado conectado: una deducción.

6. Doce repeticiones para una clave determinada reducen las posibles configuraciones del scrambler a 250: una deducción. El tráfico debería producir ese número para las claves de cada día: una inducción.

La primera ruptura del código Enigma, que Welchman duplicó, se produjo al mismo tiempo que se estaban desarrollando otros métodos para descifrarlo. Ambas cosas condujeron a la invención de varias máquinas para ayudar en el análisis que culminaron en Colossus, un precursor del ordenador digital programable. El ordenador ha conducido a su vez al desarrollo de sistemas de cifrado que son inexpugnables... hasta ahora.

CAPÍTULO 27

Sobre el modo de transmisión del cólera

En 1831 el cólera llegó por primera vez a Inglaterra. La enfermedad es terrible. Al principio, uno se siente un poco revuelto, como si hubiera comido algo que no le sentó bien. La cosa pronto va a peor. Se empieza a sufrir una fuerte diarrea, a vomitar, duelen el pecho y el estómago, y los músculos sufren espasmos. En cuestión de horas, se envejece visiblemente. El pulso se vuelve débil. La circulación es lenta. El cuerpo se enfría. Y la cara, el cuello y las manos se vuelven casi azules. Las posibilidades de morir son de una entre tres. El que muere sucumbirá en un día o dos. En Inglaterra los médicos sabían que la enfermedad venía de Asia. Era fácil de diagnosticar: la diarrea era un signo característico. Las deposiciones tenían escaso color u olor fecal. Eran acuosas y contenían pequeños granos blancos. De ahí el nombre de deposiciones de “agua de arroz”.

La enfermedad surgió en Sunderland y Newcastle-on-Tyne, puertos cercanos en la costa noreste de Inglaterra. Pero se extendió rápidamente de un lugar a otro, concentrándose en los barrios bajos. Los médicos no tenían medidas preventivas ni paliativas. Intentaban todas las posibles curas que eran capaces de imaginar, pero tenían escaso éxito. No conocían la causa de la enfermedad, su patología o cómo se extendía. Unos pensaron que era contagiosa. Otros pensaron que procedía de un miasma del aire –efluvios nocivos en la atmósfera–, quizá una consecuencia de la suciedad de los barrios bajos. Algunos pensaron que la meteorología causaba la enfermedad en quienes tenían una constitución vulnerable. Y creían que la vulnerabilidad era una consecuencia de una emoción fuerte, como el miedo. Como precaución contra los miasmas las calles de Newcastle fueron rociadas con cal. Los lugares con más riesgo se suponía que eran las zonas bajas, y la mayor parte de la ciudad estaba a una altitud segura. Sin embargo, la enfermedad se difundió incluso a los pueblos mineros de fuera de la ciudad.

El hombre que resolvió el misterio respecto al modo de transmisión del cólera

fue el Dr. John Snow, y su investigación estableció las bases para una nueva disciplina – la epidemiología, el estudio de la difusión de las enfermedades en las poblaciones–. Su teoría, que comprobó mediante ingeniosos sistemas, no fue aceptada mientras él vivió. Soportó las burlas de sus colegas y de la gente. Sus opiniones, sin embargo, fueron confirmadas después de su muerte cuando el cólera golpeó Inglaterra de nuevo en 1866, en su cuarto brote.

Lo que quiero examinar es cómo razonó Snow. Su razonamiento fue el secreto que le llevó a descubrir el mecanismo de transmisión del cólera antes de que se tuviera ningún conocimiento sobre las bacterias. La historia de la medicina es una historia del pensamiento mágico. Los doctores y los pacientes son igual de vulnerables a la falacia de post hoc ergo propter hoc (después de esto, por tanto, debido a esto). En otras palabras, los pacientes se recuperaban de las enfermedades a pesar de las sangrías, ventosas chinas, ampollas, sanguijuelas u otros curiosos regímenes con poderes curativos no demostrados, y por tanto los individuos crédulos atribuían la recuperación a esos tratamientos. El trabajo de Snow contrasta con esta historia. Su pensamiento fue un caso ejemplar de razonamiento científico.

Snow había nacido en 1813 en York, en el norte de Inglaterra. Sus padres eran pobres, pero pudieron mandarle a la escuela. Empezó su preparación médica como aprendiz de un doctor en Newcastle-on-Tyne, y estaba allí durante el primer brote de cólera. Fue enviado a un pueblo minero fuera de la ciudad, y vio por sí mismo a los mineros que subían de los pozos muriendo debido a la enfermedad. Durante sus años en Newcastle, se convirtió en vegetariano y creyente en las virtudes del agua pura: el agua potable debía ser destilada. Y se convirtió en un orador en pro de la conducta abstemia.

Después de nueve años como aprendiz de varios doctores en el noreste, consiguió ir a Londres a terminar su formación médica en la Universidad de Londres. En 1838 obtuvo los títulos de cirujano y boticario –una doble titulación que entonces era necesaria para practicar la medicina–. Se convirtió en un precursor en la utilización de éter y posteriormente cloroformo como anestésicos. Emprendió programas intensivos de investigación para examinar sus efectos sobre animales con el fin de entender la fisiología de su funcionamiento, y llegó a dominar las leyes conocidas sobre la difusión de los gases en la atmósfera. Se dio a conocer durante su vida como el doctor que en 1853 administró cloroformo a la Reina Victoria cuando estaba dando a luz a su octavo hijo –un suceso que llevó a la utilización generalizada de la anestesia en la sala

de partos–. Para entonces, había hecho una carrera de éxito como médico y anestesista en Londres. Su mayor logro, sin embargo, fue el esclarecimiento de la patología y transmisión del cólera. Murió a la edad de cuarenta y cinco años en 1858.

* * *

En el otoño de 1848, a la vez que Snow continuaba con sus investigaciones sobre la anestesia, el cólera volvió a Inglaterra en una segunda epidemia. En menos de un año, mientras continuaba el brote, él había desarrollado una teoría sobre la patología de la enfermedad y su transmisión. Muchos científicos de la era de Snow eran “inductivistas”. Creían que la investigación científica era únicamente cuestión de hacer generalizaciones inductivas a partir de las observaciones –una opinión que persistió por parte de algunos científicos casi hasta el día de hoy–. Karl Popper, el filósofo de la ciencia del siglo veinte, descartó totalmente la inducción. La ciencia, argumentaba, depende de conjeturas y refutaciones. Las conjeturas son suposiciones afortunadas y las refutaciones son observaciones que son inconsistentes con las consecuencias deductivas de las conjeturas. Una teoría científica, sin embargo, rara vez es un conjunto de suposiciones afortunadas. Ciertamente, una descripción neodarwiniana de las conjeturas es imposible. Los procesos que subyacen en la creación de una teoría científica dependen de una interacción entre inducción, abducción y deducción. En el razonamiento cotidiano se produce la misma interacción, como en mi ejemplo previo sobre los bares italianos (véase el Capítulo 13). Por consiguiente, el pensamiento científico no es especial, sino que surge de nuestras prácticas diarias. Pueden también incorporarse las probabilidades, como sucede en el análisis estadístico de datos. Los pasos críticos, sin embargo, son la abducción de hipótesis y su comprobación mediante una comparación sistemática de las consecuencias que se pueden deducir de ellas con las observaciones sobre el mundo. Como sostenía Popper, lo que diferencia a la ciencia de otros campos son las comprobaciones que podrían falsear las hipótesis.

Snow comprendió la necesidad de desarrollar teorías. Podrían basarse en parte en inducciones, pero también era necesario formular hipótesis explicativas y

deducir a partir de ellas consecuencias evaluables. Esta opinión sobre la ciencia había sido formulada anteriormente en el siglo diecinueve, y se conocía a veces como el método “hipotético- deductivo”. Muchos científicos cognitivos modernos han argumentado en favor de una concepción de la investigación científica basada en modelos. Los científicos desarrollan modelos mentales de cómo funciona el mundo, y utilizan esos modelos, tanto para crear sus descripciones de las teorías como para inferir consecuencias evaluables. La construcción mental de modelos depende de la abducción. Se apoya en un proceso de etapas múltiples, y el conocimiento, que se puede adquirir por inducción, actúa como una importante restricción. Esta descripción da sentido, como veremos, a cómo razonó Snow en la persecución de una comprensión científica del cólera.

Empezó por una serie de inducciones a partir de observaciones conocidas:

• El cólera era una única enfermedad específica.

• Se transmitía de alguna forma de una persona a otra. Su expansión seguía las rutas del comercio, y la llegada de una persona infectada a un lugar determinado llevaba con frecuencia a un brote de la enfermedad.

Razonó luego que si en algunos casos se transmitía de esta forma, lo más probable es que también lo hiciera así en todos los demás casos. Al hacer esta inducción, Snow se remite a la máxima que he citado antes: “efectos similares dependen de causas similares”. La inducción también es consecuencia de una tendencia a la parsimonia en su pensamiento. Tenía un sesgo hacia un único modo de transmisión. Este sesgo, unido a un mecanismo conocido para algunos casos y un mecanismo desconocido para otros, da lugar a la inducción de que el mecanismo desconocido es el mismo que el mecanismo conocido. Por consiguiente, su modelo inicial de la enfermedad plantea un único modo de transmisión.

A lo largo de su vida, la parsimonia de su teoría atrajo críticas. La revolución industrial en la Inglaterra del siglo diecinueve condujo a un movimiento “sanitario” en favor de la introducción de inodoros en lugar de pozos negros, de la limitación de la contaminación de las fábricas y de otras reformas sanitarias.

Sus defensores nunca pudieron aceptar la versión de Snow, y defendieron la existencia de diversos modos de transmisión de la enfermedad. Les atraía mucho más la idea de que en la atmósfera estaba presente algún contaminante general. Sin embargo, Snow fue capaz de rechazar totalmente esta teoría de los "miasmas". Su estudio de los anestésicos le había llevado a comprender las leyes de los gases, y por tanto dedujo que la distribución de las partículas nocivas en el aire sería demasiado difusa como para transmitir la enfermedad. Como escribió más tarde: "A medida que los gases expulsados por las sustancias en putrefacción se difunden en el aire, la cantidad en un espacio dado es inversamente proporcional al cuadrado de la distancia respecto a la fuente. Por ello, un hombre que trabaja con el rostro a una yarda [unos 90 cm] de las sustancias pestilentes respiraría diez mil veces más cantidad del gas expulsado que una persona que viva a cien yardas de ese punto."

Una opinión similar, sin embargo, era que la enfermedad se transmite de una persona a otra por la inhalación de algo contagioso del aire que haya emanado de las víctimas. Snow había partido de esta opinión, pero, una vez que inició sus propias investigaciones sobre la enfermedad, razonó que el mecanismo era improbable. El veneno variaría su concentración en el aire. Algunas personas tendrían que resistir a altas densidades, porque estaban en contacto con las víctimas pero no sucumbían, mientras que otras tendrían que ser vulnerables a densidades muy bajas, porque estaban lejos de cualquiera que tuviera la enfermedad y sin embargo sucumbían. Era improbable que la vulnerabilidad al veneno pudiera variar tan ampliamente.

En sus informes Snow decía que para desarrollar una explicación de la patología del cólera estableció una analogía con otras enfermedades. Los médicos sabían que los óvulos de ciertas lombrices intestinales, si se ingerían, podían infectar el tubo digestivo. Snow hizo una inferencia abductiva por analogía con este mecanismo: de manera similar, desconocidas partículas venenosas del cólera podrían ingerirse y transmitir la enfermedad al intestino. Estas partículas podrían ser organismos vivos o no, pero estaban organizadas y obedecían a las leyes de la química. Hizo una abducción más. Las partículas se multiplicaban en el intestino, e inflamaban su membrana mucosa de manera que expulsaba las sales y la parte acuosa de la sangre, que los pacientes evacuaban después. Esta hipótesis explicaba dos viejas observaciones: los pacientes se deshidrataban y su sangre se espesaba. Este espesamiento, infirió, también explicaba los problemas circulatorios de los pacientes y la dificultad para respirar.

¿Cómo pasaban de una persona a otra las hipotéticas partículas de cólera? Cuando la enfermedad se transmitía dentro de una casa, una abducción verosímil era que la diarrea de las víctimas infectaba su ropa y la ropa de cama. Su falta de color y olor fecales podría llevar a otros a recoger esta ropa. Si luego preparaban la comida sin lavarse las manos, podrían pasar entonces las partículas a la comida y la bebida. Ellos y otros ingerirían después las partículas.

Pero la enfermedad podía saltar también a grandes distancias. Puede que la vieja preocupación de Snow por el agua potable purificada le llevara a abducir que las partículas de cólera podrían entrar en el suministro de agua potable. Si aguas residuales infectadas contaminaban el suministro de agua, entonces la enfermedad podría transmitirse en una amplia área.

La construcción de la teoría se basó en el conocimiento general y la experiencia particular de Snow sobre temas como las leyes de los gases. Como en cualquier teoría científica, se guio por ciertos resultados establecidos. Estaba preparado para hacer inducciones que simplificaran las observaciones que había que explicar: infirió que el modo de transmisión era único en lugar de múltiple. Utilizó una analogía para abducir una hipótesis clave: las partículas de cólera infectaban el intestino de la misma forma en que podían infectarlo los óvulos de las lombrices. La inflamación llevaba a que se liberaran dentro del intestino las sales y la parte acuosa de la sangre. Estos productos se evacuaban mediante la diarrea de las víctimas.

Dada esta patología, la transmisión de la enfermedad exigía una causa por contacto. Una vez que había desechado la teoría de los miasmas, quedaba solamente la inhalación de las partículas nocivas o su ingestión. La inhalación no conseguía explicar la difusión de la enfermedad de un lugar a otro a menos que uno estuviera dispuesto a hacer una asunción improbable respecto a la vulnerabilidad de los individuos en relación con la enfermedad. La ingestión era probable dentro de una casa, y por tanto lo único que faltaba por explicar era la transmisión a largas distancias. Admitida la restricción de que la materia fecal era la transmisora y el intestino era el receptor, el agua potable era un destacado canal de comunicación. Ciertamente, es difícil concebir ninguna otra posibilidad.

La teoría de Snow establecía una cadena de causación por contacto físico que explicaba tanto la transmisión de la enfermedad como su patología:

1. las excreciones de las víctimas de cólera contenían partículas venenosas;

2. estas partículas podrían transmitirse a otros por contacto físico con ropas manchadas y podrían entrar en las aguas residuales;

3. en el último caso, si las aguas residuales contaminaban el suministro de agua, entonces otras personas que bebieran el agua ingerirían las partículas;

4. cuando las partículas fueran ingeridas, se multiplicarían e inflamarían el intestino causando la enfermedad en las víctimas infectadas;

5. la inflamación causaba la extracción del suero de la sangre, es decir, las sales y su parte acuosa;

6. esta extracción causaba que la sangre se espesara, y esto causaba problemas secundarios en la circulación (el aspecto azul de los pacientes) y en la respiración, que podrían dar lugar a la muerte;

7. las sales y la parte acuosa de la sangre, junto con las partículas venenosas de la enfermedad, entraban en el aparato digestivo.

Y así continuaba el ciclo de infección.

La teoría causal descansaba tanto en la abducción como en inducciones a partir de las observaciones. Solo le faltaba un componente: la naturaleza de las partículas venenosas. Pero implicaba que un pequeño número de ellas fuera suficiente para causar la enfermedad, y por tanto fueran capaces de multiplicarse. Implicaba también que mantenían su potencia y no se diluían en grandes cantidades de agua. Snow infirió que eran entidades organizadas, quizá incluso organismos vivos.

* * *

La teoría era verosímil, y Snow dedujo algunas consecuencias comprobables. La

más destacada era que si el agua potable era el canal de transmisión entonces el cólera debería tender a aparecer en cualquier sitio en que el suministro de agua pudiera haber sido contaminado por aguas residuales que contuvieran partículas venenosas. Su búsqueda de pruebas en relación con este punto le llevó a convertirse en un pionero de la epidemiología, y después de la epidemia de cólera de 1848-49 ayudó a fundar la Sociedad Epidemiológica de Londres.

En principio, Snow reunió las pruebas de las que iba teniendo conocimiento. Describió dos informes que había leído sobre la epidemia de cólera de Londres de 1848. El primero describía un brote en Horsleydown. En un bloque de edificios la enfermedad había sido incontrolada; en otro bloque cercano había habido un único caso. La única diferencia entre ambos bloques, observaba el autor del informe, era que el agua potable del edificio infectado se extraía de un pozo al mismo nivel que un sumidero de aguas residuales. Un segundo informe describía una diferencia similar en Albion Terrace en Wandsworth. No parece que Snow utilizara mapas que apoyaran sus reflexiones, aunque sí los utilizaba para ayudar al público a extraer de ellos consecuencias por deducción. Sin embargo, a menudo imaginaba las disposiciones espaciales de los sumideros y las tuberías de agua. En Albion Terrace, los sumideros de ladrillo y los pozos negros estaban más altos que la tubería que suministraba el agua potable. Una fuerte tormenta de agua había provocado que se desbordaran los sumideros, y dos días más tarde se produjo el primer caso de cólera. El desbordamiento podría haber contaminado el suministro de agua. Debido a que no se produjo ningún otro brote en la región, este caso era un factor en contra de la hipótesis de que la causa de la enfermedad era un miasma. Por consiguiente, la no existencia de la enfermedad era contraria a una consecuencia deducida de la teoría de los miasmas, y actuaba como un contraejemplo para esa teoría.

Su siguiente paso fue recoger datos sistemáticos. Elaboró una tabla de muertes por cólera en las principales zonas de Londres. Al sur del río Támesis el índice de mortalidad por cólera era siete veces mayor que al norte del río. El sur de Londres obtenía el suministro de agua de compañías que extraían el agua de la región del Támesis afectada por las mareas, más abajo que los principales desagües de aguas residuales de la ciudad. Por consiguiente, el agua podría estar contaminada por partículas venenosas de cólera. El norte de Londres obtenía su suministro de agua de compañías que extraían el agua de la zona del Támesis por encima de los desagües de aguas residuales. Era mucho menos probable que su agua estuviera contaminada. La diferencia daba que pensar, pero no era decisiva. Snow continuó coleccionando pruebas en favor de esta teoría después de la

publicación de su monografía sobre la enfermedad. Mantuvo correspondencia con muchas personas de todo el país para reunir datos, tanto sobre la primera epidemia de 1831 como sobre la más reciente de 1848. Describió casos en los que se produjeron brotes y casos en los que no se habían producido pero, según las teorías de sus rivales, deberían haberlo hecho. La diferencia clave, argumentaba, era si las aguas residuales podrían haber contaminado o no el suministro de agua.

Quizá la mejor de sus pruebas fue un experimento natural que afectaba a dos ciudades inglesas: Exeter y Hull. Exeter padeció la epidemia de 1831, pero luego su planta de suministro de agua se había trasladado aguas arriba de la salida de aguas residuales. Tuvo muchas menos víctimas en la epidemia de 1848. El sino de Hull fue el contrario: pocas víctimas en la primera epidemia; su planta de suministro de agua se trasladó a una zona contaminada del río; y tuvo muchas más víctimas en la segunda epidemia. A pesar de ello, los defensores de las teorías alternativas no se convencieron, y continuaron apoyando la teoría de los miasmas.

Lo que faltaba en el análisis de Snow era un estudio general de suministro de agua y cólera. Cuando se desató una tercera epidemia en el verano de 1853, Snow fue capaz de llevar a cabo un “gran experimento”. Debido a que para la hipótesis de los miasmas la altura sobre el nivel del mar era importante, este experimento crucial tenía que contraponer la naturaleza del suministro de agua con la altura sobre el nivel del mar, a fin de determinar qué variable predecía el índice de víctimas del cólera. Snow se dio cuenta de un rasgo interesante en las cifras oficiales de muertos por cólera. El abastecimiento de agua en distintas áreas de Londres lo realizaban dos compañías rivales, una tomaba su agua por encima de la zona de mareas del Támesis mientras que la otra no. Analizó los respectivos índices de mortalidad. Nadie había muerto de cólera en aquellos subdistritos abastecidos solo por la compañía que extraía el agua en la zona alta, pero la mortalidad en los subdistritos abastecidos solo por la compañía que extraía el agua en la zona baja era del 114 por 10.000. Sin embargo, había un inconveniente. Los subdistritos tenían diferentes características. Los saludables eran suburbios residenciales. Los insalubres eran zonas superpobladas y sumidas en la pobreza que pertenecían a las zonas deprimidas de la ciudad.

La comparación crucial solo se podría hacer en los subdistritos abastecidos por ambas compañías. En estos, las tuberías de ambas recorrían todas las calles, y por tanto una casa podía recibir agua de una compañía mientras que la casa

contigua la recibía de la compañía rival. Todo lo que se necesitaba era determinar los índices de mortalidad en ambos tipos de casas. Snow decidió centrarse en las muertes por cólera, y determinar luego qué compañía abastecía de agua las casas de las víctimas. El cólera había decrecido hacia finales de 1853, pero creció de nuevo fuertemente en el siguiente verano. Snow dejó la práctica médica con el fin de poder recorrer las calles de Kennington en el sur de Londres para reunir datos relevantes. De cuarenta y cuatro muertes producidas allí durante las primeras cinco semanas de la renovada epidemia, treinta y ocho se produjeron en casas abastecidas por la compañía que extraía el agua de la zona baja. Es absolutamente improbable que esta diferencia se produzca por casualidad. Si yo lanzara al aire una moneda cuarenta y cuatro veces, y saliera cara treinta y ocho, usted tendría motivos para pensar que le estaba engañando. Las pruebas estadísticas no se habían inventado en la época de Snow pero, si se lanza honradamente una moneda sin trucar, las posibilidades reales de que se produzca ese resultado, u otro más extremo, son que debería ocurrir menos de cinco veces en diez millones de experimentos en que se lanzara la moneda cuarenta y cuatro veces.

Snow continuó reuniendo pruebas con un colega médico. No era fácil averiguar qué compañía suministraba el agua en muchas casas, porque a menudo las facturas de agua las pagaban arrendadores ausentes. Pero la experiencia química de Snow le echó una mano. Fue capaz de diseñar una sencilla prueba química que distinguía entre el agua de las dos compañías, porque se diferenciaban en cuanto a salinidad. Él y su colega descubrieron que se produjeron 286 muertes por cólera en las casas abastecidas por la compañía que extraía el agua en la zona baja, mientras que solo se produjeron catorce muertes por cólera en las casas abastecidas por la compañía que extraía el agua en la zona alta. ¡Imagínese obtener ese resultado al lanzar una moneda!

A medida que continuara la epidemia, la diferencia debería disminuir porque entonces los individuos corrían el riesgo de contraer la enfermedad a partir de otras personas, en lugar de a partir del agua contaminada. De ahí que Snow predijera que la diferencia en los índices de mortalidad entre las casas suministradas por las dos compañías debería disminuir. También corroboró esta predicción. Consiguió persuadir a los encargados del registro civil de que recogieran datos para él. Estos datos necesariamente iban a ser menos fiables que sus propios resultados. Los empleados no podían llevar a cabo la prueba para determinar qué compañía abastecía a una casa si los ocupantes no lo sabían. Sin embargo, las cifras mostraron una gran y parecida diferencia en cuanto al

riesgo entre las dos compañías, siendo la relación de muertes de aproximadamente cinco a una. Snow publicó estos datos en la segunda edición de su libro en 1855. La ironía era que la mejora del alcantarillado de Londres –gracias a los sanitarios– había llevado a que se vertieran más aguas residuales al Támesis, y como resultado a un aumento en las muertes por cólera.

Cuando Snow hubo recogido todos estos datos, llevó a cabo una comprobación matemática de su teoría. Su asunción básica era que el factor más importante en la difusión de la enfermedad era el suministro de agua, y por tanto el número de muertes debería correlacionar con el número de casas que recibían agua contaminada. Examinó treinta y un subdistritos de Londres. Había habido 160 muertes por cada 10.000 casas abastecidas por la compañía que extraía el agua en la zona baja, y veintisiete muertes por cada 10.000 casas abastecidas por la compañía que extraía el agua en la zona alta. Su teoría predecía que estas proporciones deberían producirse también en cada uno de los subdistritos, y examinó la diferencia entre ellas y las cifras reales. Describió el resultado como "una relación muy próxima".

* * *

En contraste con la lenta pero decisiva acumulación de pruebas estadísticas, un único brote de cólera en Broad Street (ahora Broadwick Street) en el Soho del centro de Londres fue el episodio más gráfico de las investigaciones de Snow sobre el cólera. Cuando se enteró del brote, que se produjo a solo unos minutos andando desde su casa, sospechó inmediatamente de la contaminación del agua de un surtidor. Recordó un incidente similar en la epidemia de 1848. Como escribió más tarde, fue a Golden Square, en el Soho, con esta hipótesis en mente. Inspeccionó el agua de un surtidor en Broad Street y, tanto él como un experto en observaciones microscópicas, detectaron en ella materia orgánica. Siguiendo con sus pesquisas, un informador le dijo que el agua tenía un olor desagradable inmediatamente antes de que empezara el brote. Pero de nuevo la prueba daba que pensar pero no suponía nada más.

Snow fue capaz de inferir entonces otra forma de comprobar su hipótesis. Los individuos que vivieran más cerca de ese surtidor que de ningún otro deberían

haber sucumbido a la enfermedad con más frecuencia que aquellos que vivieran más cerca de otro surtidor. Calculó el punto que en cada calle equidistaba del surtidor de Broad Street y de otro surtidor. Sus pesquisas casa por casa mostraron que la mayoría de las víctimas habían vivido en casas más cercanas al surtidor de Broad Street que a cualquier otro. Snow se presentó en una reunión de las autoridades locales responsables de la salud, y les persuadió de que quitaran la manivela del surtidor. La epidemia se estaba reduciendo, pero después de quitar la manivela no se produjo ningún nuevo caso.

Snow continuó investigando el brote. Descubrió que un taller de la zona se había librado en gran medida: tenía su propio pozo. Se asoció al párroco local, que le ayudó a encontrar a miembros de las familias de las víctimas que se habían marchado de la zona. La inmensa mayoría de las víctimas había bebido agua del surtidor. El párroco localizó también el primer caso: una bebé de cinco meses que había muerto de “diarrea” en una de las casas cercanas al surtidor. Mientras yacía agonizante, otros habitantes de la zona empezaron a caer enfermos. Una inspección cuidadosa del surtidor reveló que una cloaca cercana podría haber contaminado el pozo que le aportaba el agua.

Durante el curso de sus investigaciones, Snow preparó dos mapas del área, mostrando el número de muertes en cada casa. El primer mapa, que publicó en la segunda edición de On the Mode of Communication of Cholera, localiza el surtidor de Broad Street en el lugar equivocado. El segundo mapa corrige este error, y también contiene un rasgo innovador. Snow dibujó una línea alrededor del área que delimitaban los puntos equidistantes entre el surtidor de Broad Street y algún otro surtidor. La Figura 27.1 presenta este mapa. Snow lo utilizó para hacer la sencilla deducción de que las muertes se reducían casi a cero fuera del perímetro de la línea –como ayuda en una investigación oficial sobre el brote.

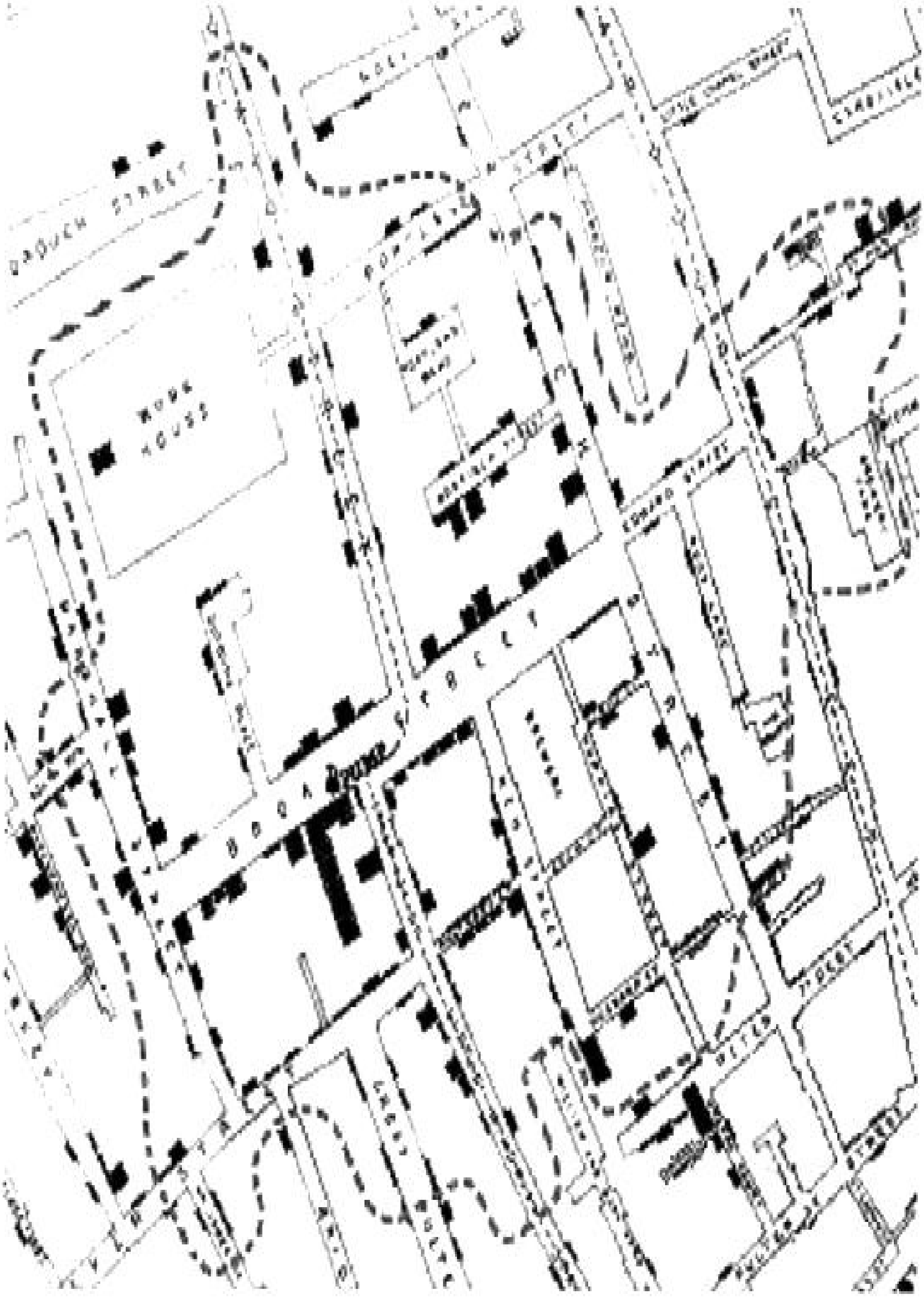

WORK HOUSE
BROAD STREET

Fig. 27.1. Segundo mapa de John Snow sobre el brote de cólera en Broad Street. El surtidor está cerca del centro de Broad Street. La línea discontinua, que he realzado además de hacer algún otro arreglo electrónico a partir de una reproducción del mapa, muestra la línea de puntos equidistantes entre el surtidor de Broad Street y los otros surtidores trazada por Snow. Cada rectángulo negro representa a una persona muerta durante el brote.

* * *

Snow se dio cuenta de que su teoría implicaba medidas prácticas para prevenir la difusión de la enfermedad. Era necesario bloquear el camino fecal-oral. Una deducción elemental era que los individuos en contacto con las víctimas necesitaban asegurarse de que no ingerían partículas del cólera. Podría ser efectiva una cuarentena de las víctimas, pero, al contrario que otras teorías, la versión de Snow no requería que los individuos infectados fueran aislados. Se necesitaba higiene, no cuarentena. Si los cuidadores habían tocado a personas infectadas o sus ropas de cama, necesitaban ser escrupulosos en el lavado de manos. Si el canal de transmisión era el suministro de agua, entonces la gente debería hervir el agua antes de beberla, y evitar beber ninguna que pudiera estar contaminada. Las compañías de agua deberían extraerla en zonas anteriores a cualquier vertido de aguas residuales, y comprobar que las cañerías que distribuían el agua no podían ser contaminadas por filtraciones de alcantarillas o pozos negros. Lo mismo que su teoría, estas propuestas no fueron aceptadas mientras él vivió.

* * *

El esclarecimiento del cólera por Snow fue un caso ejemplar de razonamiento científico. Tuvo a su disposición algunos hechos, que dieron lugar a algunas

inducciones sobre la epidemia, como por ejemplo que el cólera era una única enfermedad específica. Estas inducciones dependían del conocimiento: los médicos habían descubierto mucho antes la existencia de enfermedades específicas. A partir de este modelo de enfermedad, Snow formuló una teoría parsimoniosa de su patología, que explicaba sus efectos causales sobre aquellos a los que infectaba. De manera similar, mediante un proceso de abducción, explicó la transmisión de la enfermedad. En sus reflexiones fue fundamental una analogía entre la transmisión conocida de las lombrices y la transmisión desconocida de las hipotéticas partículas de cólera. En 1854, sin que Snow lo supiera, el italiano Filippo Piccini identificó el Vibrio Cholerae, la bacteria que causa la enfermedad. El mundo prestó poca atención a su trabajo. Su descubrimiento se le sigue atribuyendo a Robert Koch, que redescubrió la bacteria. La conexión final en el pensamiento de Snow fue una abducción respecto a la transmisión de las partículas. Dada la asunción de la causación por contacto, un modelo espacial sencillo da lugar a dos caminos: uno de corto alcance, vía contacto con ropas contaminadas, y uno indirecto y de largo alcance. La patología de la enfermedad apoyaba la hipótesis de la contaminación del suministro de agua.

Una vez que Snow tuvo su teoría, dedujo a partir de ella varias consecuencias comprobables. Estas consecuencias implicaban la necesidad de ciertos tipos de datos, que Snow recogió. Para alguien imparcial, estas consecuencias proporcionaron argumentos decisivos en apoyo de su teoría y en contra de las que rivalizaban con ella. Quienes proponían estas explicaciones, que en su mayoría pertenecían a la profesión médica y sanitaria, no se convencieron y no fue hasta después de la muerte de Snow y la llegada de la cuarta epidemia de cólera en 1866 cuando los profesionales y la población empezaron a aceptar sus ideas.

En el sofocante verano de 1858, mientras John Snow yacía moribundo, el río Támesis empezó a emitir un terrible olor. Él murió de un derrame cerebral el 16 de junio. Dos días más tarde, el periódico Times describía la “gran fetidez”, informando de que había hecho desalojar a los miembros del Parlamento y exigiendo leyes para enfrentarse a ella. En el Parlamento se colgaron cortinas empapadas en cal como protección. Pero, en contra de los partidarios de los miasmas, no se produjeron enfermedades epidémicas. Había empezado la reivindicación de Snow. Hay algunas cuestiones irónicas para terminar. Al ardiente abstemio se le conmemora actualmente en Broadwick Street con un pub, The John Snow, y con una réplica del surtidor de triste memoria situada, por

supuesto, en el lugar equivocado.

CAPÍTULO 28

Cómo razonamos

Nuestra capacidad para razonar es vital. Cuanto mejor razonamos mejor es nuestra vida. Estamos más sanos, vivimos más, y tenemos más éxito en lo académico y en la calle. En la vida real, hacemos muchas inferencias sencillas, y las hacemos casi sin darnos cuenta de que estamos razonando. Ya describí lo que sucedió la primera vez que entré en un bar italiano a tomar un capuchino. Tuve que inferir que primero se paga y luego se pide en la barra. Esta inducción estaba apoyada por una deducción, que el bar no era excepcional, porque si lo hubiera sido entonces otros clientes tendrían que haber hecho el mismo recorrido que yo – a la barra, a la caja registradora y de vuelta a la barra–. Hubo otro tipo de inferencia que reforzó mi inducción. Hice una abducción que explicaba el modus operandi del bar. El lugar estaba demasiado concurrido como para que los hombres de la barra manejaran una caja registradora. Una solución eficiente era que los clientes presentaran recibos, en los que mostraran aquello por lo que habían pagado, y que los camareros prepararan las correspondientes bebidas.

¿Cómo razonamos? Depende de cómo se inicia una inferencia. Algunas inferencias se inician inconscientemente pero dan lugar a conclusiones de las que somos conscientes. Lo llamamos conjeturas, presentimientos o intuiciones. Son rápidas y a menudo exactas, pero solo hacen frente a propiedades sencillas o relaciones sencillas. Miramos las fotografías de dos personas e inmediatamente nos formamos la impresión de que una de ellas parece más competente que la otra (véase la Figura 4.1). No podemos expresar las razones para esta inferencia, pero es sólida y la mayoría de nosotros la hacemos. Hacemos otras inferencias que son totalmente inconscientes –ni siquiera somos conscientes de haber llegado a ningún tipo de conclusión– pero, como muestran los experimentos, estas inferencias se traslucen en cambios en nuestra conducta. Saber exactamente qué parte de nuestra conducta tiene estas raíces inconscientes es una cuestión no resuelta, pero existe un límite. Para fastidio de Madison Avenue, la publicidad subliminal no nos hace comprar ni siquiera refrescos, menos aún vehículos todoterreno. Los procesos inconscientes sin acceso a la memoria de

trabajo no son capaces de hacer las inferencias complejas que son necesarias para esas operaciones.

En Inglaterra, si hubieran rehusado servirme en un bar, me habría enfadado. Pero, para cuando entré en el bar, ya había hecho la inducción intuitiva de que los italianos eran amigables, y por tanto no inferí que el camarero estaba siendo hostil, y no me enfadé. Las inferencias pueden llevar a emociones, pero no somos conscientes de la transición del pensamiento al sentimiento. Ciertamente, no somos conscientes de ningún proceso mental, ni siquiera de aquellos que construyen el mundo en la mente consciente.

Una opinión influyente sobre las enfermedades mentales, como fobias o compulsiones obsesivas, es que están causadas por un razonamiento defectuoso. La terapia "cognitiva" se basa en esa asunción. Pero, nadie parece haber examinado esa asunción con anterioridad a los estudios a los que me refiero en la Parte II del libro. Comparan a participantes con tendencia a diversos tipos de neurosis con grupos de control sin esa tendencia. Lo que los resultados revelaban era que aquellos que tienen tendencias neuróticas son mejores en el razonamiento deductivo que los participantes de control, pero solo cuando razonan sobre temas relevantes para su neurosis. La diferencia desaparece con otros temas, aunque los que tienen tendencia a la neurosis no tienen peores resultados que los participantes de control.

Si el razonamiento defectuoso no causa los signos y síntomas neuróticos, ¿qué los causa? Una posibilidad son las reacciones emocionales anormales. Hacemos una transición inconsciente a una emoción que es adecuada para lo que sabemos de la situación pero que tiene una intensidad desproporcionada. Un individuo sin historial de enfermedad psicológica entró en una "tienda de sudar" [sweat longe] india: estaba totalmente oscuro, e inmediatamente le invadió una ansiedad tan intensa que se vio obligado a salir. Una vez fuera, razonó para sus adentros que su temor era injustificado, y que se estaba volviendo loco. Volvió adentro –solamente para sentirse invadido de nuevo por la misma intensa ansiedad, y por tanto tuvo que volver a salir–. ¿Qué causaba la ansiedad? No podía decirlo. En el pasado, este tipo de vacío intelectual se ha llenado de diversas formas. Los psicoanalistas lo han llenado con apelaciones a deseos reprimidos alejados del objeto de nuestras emociones. Los terapeutas cognitivos lo han llenado con pensamientos a los que podemos acceder si persisten en interrogarnos. Pero la teoría de la hiperemoción, como yo la llamo, implica que la evaluación de la situación es sensata –los lugares oscuros pueden ser peligrosos– y que lo que ha

ido mal es solo la intensidad de la emoción. Algunos somos propensos a estos episodios como resultado de factores genéticos, y la propensión se puede exacerbar en entornos estresantes. Cuando tenemos una emoción aberrante, razonamos sobre ella. El razonamiento nos centra en el problema, y de ese modo intensifica nuestra emoción y generaliza las situaciones que la precipitan. Si esta teoría es correcta entonces lo que la terapia necesita es reeducar el sistema de transiciones inconscientes que lleva del conocimiento a la emoción.

* * *

¿Cómo razonamos? Los filósofos de la Ilustración sostenían que seguimos las "leyes del pensamiento". Estas leyes se hacen explícitas mediante la lógica formal y el cálculo de probabilidades. En lógica, las leyes del pensamiento se remontan a los griegos, e incluyen la ley de no contradicción: nada es ambas cosas A y no-A, donde A es cualquier propiedad en la que podamos pensar. Las leyes se plasman en las reglas formales de inferencia de la lógica, y podemos aprender a utilizar estas reglas para hacer inferencias. La cuestión es si, al razonar, quienes no tienen formación lógica se apoyan en la forma o en el contenido. Entendemos el significado y la significación de aquello sobre lo que razonamos. Por consiguiente, podemos ser escépticos respecto a que utilicemos reglas de inferencia que se apliquen forzosamente a cualquier premisa que tenga la forma lógica apropiada. El proceso exigiría que recuperáramos la forma y nos deshiciéramos del contenido, como un político que aprendiera a articular respuestas perfectas a las preguntas de los periodistas sin entender el significado de lo que decía. Nuestras intuiciones sobre los procesos mentales no tienen mucho valor, pero los psicólogos pueden ir más allá de sus intuiciones para realizar experimentos. Y he informado de muchos resultados que implican que no razonamos en base a la forma lógica.

Una observación es la extraordinaria dificultad que tenemos, tanto para expresar como para analizar argumentos en un formato lógico. Como mostré, personas diferentes tienden a hacer análisis lógicos diferentes del mismo argumento. Es una dificultad extraña si nuestro razonamiento es formal. Lo natural debería ser construir y deconstruir argumentos en términos formales.

Otra observación es que las propiedades inferenciales de las conectivas, las relaciones y los cuantificadores, son imposibles de captar de una manera sencilla. Utilizamos nuestro conocimiento del significado, la referencia y el mundo para modular su interpretación. Por consiguiente, ninguna conectiva proposicional del lenguaje cotidiano, como por ejemplo “si” u “o” puede ser considerada de la misma forma que en lógica. Por ejemplo, el que una conjunción como “Él se cayó de la bicicleta y se rompió la pierna” sea cierta depende de algo más que del hecho de que las dos cláusulas que la forman sean ciertas: para que la proposición sea cierta los sucesos deben producirse también en el orden temporal correcto. De la misma forma, una disyunción como, “O bien él está en Río o en Brasil”, provoca una interpretación según la cual él está en Brasil y puede, o no, estar en Río. El mismo problema se transmite al intento de captar en axiomas las propiedades lógicas de términos relacionales, tales como por ejemplo “más alto que”. Estas propiedades dependen tanto del contexto, que la tarea es casi imposible. Si nos enteramos de que la Reina de Corazones era más alta que el Conejo Blanco, que a su vez era más alto que Alicia, podemos inferir que la Reina es más alta que Alicia. Pero todo cambia si la premisa continúa: hasta que Alicia bebió la poción mágica. De nuevo, uno se pregunta por qué, si el razonamiento se basa en la forma, el lenguaje es tan pródigo en salirse de la formalidad.

Sin embargo, otra observación es que las personas cometen errores sistemáticos al razonar en la vida real. Como ya mencioné, los ingenieros a cargo de la central nuclear de Chernobil siguieron adelante con su desastroso experimento. Sabían que para que el experimento continuara las turbinas tenían que estar girando lo suficientemente rápido como para generar electricidad de emergencia, y sabían también que las turbinas no estaban girando lo suficientemente rápido como para hacerlo. Deberían haber inferido que no deberían continuar con el experimento. Pero, como han mostrado estudios psicológicos, la inferencia no es fácil.

Las reglas formales de inferencia no permiten errores sistemáticos. Dan lugar a inferencias válidas, y por tanto, como ha dicho un importante defensor del método, cualquier error que se produzca debería ser esporádico y fortuito –resultado de un “accidente” en la mente más que consecuencia de sus principios básicos–. En realidad, los experimentos muestran que sí cometemos errores sistemáticos en nuestro razonamiento. Y hay grandes diferencias en la capacidad de razonar entre una persona y otra. En muchos experimentos, esas diferencias no podrían ser mayores. Esto es una mala noticia para las leyes del pensamiento. Si las leyes están en nuestra mente, entonces deberían formar parte de nuestro

conocimiento innato porque la mayoría de nosotros no recibe formación en relación con ellas. Por tanto, todos deberíamos estar dotados de ellas, lo mismo que todos estamos dotados de los medios para adquirir un lenguaje natural. Y por tanto, todos deberíamos adquirir una capacidad comparable en cuanto al razonamiento, lo mismo que hacemos al hablar nuestra lengua materna, la cual adquirimos sin que nos la enseñen.

Quizá en nuestra mente no hay leyes de pensamiento si no las hemos estudiado en la obra de George Boole y otros lógicos, o bien hemos tenido una oportunidad de desarrollar algunas generalizaciones formales como resultado de la experiencia realizando experimentos sobre razonamiento. Podría usted suponer –como han hecho los psicólogos evolucionistas, que, en lugar de eso, el cerebro ha desarrollado un amplio conjunto de herramientas intelectuales diferentes– como una de esas navajas que tiene veinte artilugios, desde una lima de uñas a un instrumento para quitar piedras de las pezuñas de los caballos (un "mondapezuñas"). Pero eso tampoco parece correcto, porque nuestros mecanismos inferenciales parecen coherentes tanto si estamos razonando sobre bares como sobre el cólera, y no parece que nos topemos con principios contradictorios cuando razonamos sobre el riesgo de coger el cólera de alguien que nos ha engañado. Los conflictos que se producen en nuestro razonamiento son entre nuestras intuiciones inconscientes y nuestras inferencias conscientes.

* * *

Si no nos basamos en la forma ni en navajas suizas intelectuales, entonces, ¿cómo razonamos? He llegado a la teoría de que nos basamos en el contenido. Utilizamos la percepción, los significados de las palabras y las oraciones, la significación de las proposiciones que estas expresan y nuestro conocimiento. Ciertamente, utilizamos todo lo que tenemos para imaginar posibilidades, y representamos cada posibilidad en un modelo mental del mundo. En el caso ideal, debido a lo rápidamente que disminuye nuestra capacidad para enfrentarnos a posibilidades múltiples, tendríamos que imaginar solo una única posibilidad. El uso de modelos es inabarcable. Cuantos más modelos tenemos que tener presentes, peor resulta nuestro rendimiento –hasta el punto en que ningún instrumento finito, como por ejemplo la mente humana o un ordenador

digital, pueden enfrentarse con una inferencia en un tiempo viable.

Un modelo no es una imagen, sino la estructura abstracta que subyace en las imágenes y que representa contenidos que no podemos visualizar. Pero, como una imagen, un modelo es “icónico”, utilizando el término de Peirce para una representación cuya estructura corresponde a la estructura del mundo –al menos a la que extraemos a partir de la percepción–. Un modelo icónico tiene la ventaja de que puede dar lugar a relaciones nuevas entre las entidades que representa. Estas relaciones emergentes son el elemento básico de las conclusiones. Sin embargo, no todo puede representarse como un icono. La negación, las posibilidades implícitas y el estatus de un modelo –el hecho de que represente, pongamos por caso, un hecho, una creencia, una posibilidad o incluso algo que no sucedió– depende de anotaciones proposicionales que funcionan conjuntamente con el sistema interpretativo.

Las conclusiones que extraemos dependen de los modelos en los cuales se cumplen. Inferimos que una conclusión tiene que ser cierta si se cumple en todos los modelos de nuestras premisas, probable si se cumple en la mayoría de ellos y posible si al menos se cumple en uno de ellos. Y, si consideramos la proporción de modelos en los que una conclusión se cumple, o representamos qué oportunidades tiene de aparecer, podemos inferir que una conclusión tiene una cierta probabilidad. A menudo, sin embargo, vamos más allá de la información que se nos da a fin de hacer una inducción, o introducimos nuevos conceptos, además de los que se incluyen en las premisas, a fin de explicar un hecho –hacemos una abducción–. En todos estos casos, en el centro de nuestro razonamiento están los modelos mentales de las posibilidades.

* * *

La teoría afronta dos retos que van en direcciones opuestas. Debe dar cuenta de la naturaleza de los errores que cometemos al razonar y debe explicar nuestro potencial de racionalidad. La teoría de modelos predice diversas fuentes para el error humano, y ya he descrito experimentos que las corroboran. Una poderosa fuente de error es la necesidad de considerar múltiples posibilidades. Este punto es fácil de establecer en los experimentos: cuantas más posibilidades tenemos

que considerar, más difícil es la tarea –nos lleva más tiempo y cometemos más errores–. Ciertamente, cuantas más son las posibilidades sobre las que tenemos que pensar, más confusos solemos estar. El mismo problema se produce en la vida real. Cuando en 1979 se desconectó la turbina en la central nuclear de Three Mile Island, a los operadores les llevó un par de horas averiguar cuál de las muchas posibilidades existentes era responsable de ello.

¿Cuántos modelos son demasiados para que los captemos? Los estudios han mostrado que solamente con que tengamos que tener presentes tres posibilidades nuevas pero distintas, ya tenemos problemas. Si duda sobre este punto, pregúntese a sí mismo qué se sigue de estas dos premisas:

O bien se puede tener un bonus o bien se puede trabajar menos horas, pero no ambas cosas.

O bien se puede trabajar menos horas o bien se puede tener un incremento de pensión, pero no ambas cosas.

Cada premisa da lugar a dos posibilidades, pero hay que combinarlas, y eso no es tan fácil. “O bien” significa que una proposición es verdadera y la otra es falsa, pero no se sabe cuál es cuál. Por tanto, ¿cuáles son las posibilidades? En realidad solo hay dos:

Se puede tener un bonus y un incremento de pensión o bien se puede trabajar menos horas.

La inferencia puede con la mayoría de nosotros.

Un programa de ordenador que simula la teoría de modelos predice nuestros más persistentes errores. Son los diversos tipos de inferencia ilusoria que he descrito

a lo largo del libro. Como ejemplo final, consideremos un aviso basado en la página web de un profesor de química respecto a un curso universitario:

O bien se obtiene una nota de “aprobado” si no se hace el examen final, o bien si se hace se obtiene una nota numérica.

Parece razonable, ¿no? Sin embargo, no dice lo que el profesor quiere decir. Como señalé hace un momento, el conectivo “o bien” quiere decir que una proposición es verdadera y la otra es falsa. El profesor no quiere decir eso. Él quiere decir que ambas proposiciones son ciertas, pero que se aplican en circunstancias diferentes.

Por consiguiente, lo que debería haber dicho es:

Se obtiene una nota de “aprobado” si no se hace el examen final, sin embargo, si se hace, entonces se obtiene una nota numérica.

Por tanto, siempre que tenga dudas sobre si decir “o bien” o decir “y”, inténtelo diciendo “sin embargo”. Su interpretación lógica es similar a “y”, pero tiene en cuenta las situaciones contrapuestas. Las ilusiones surgen de nuestro sistema de razonamiento. Debido a que no podemos tener presentes más que unas pocas posibilidades, tendemos a centrarnos en lo que es verdadero a expensas de lo que es falso. La misma desviación puede subyacer en la dificultad de la inferencia de Chernobil que parte del condicional, “Si el experimento tiene que continuar entonces la turbina tiene que estar girando lo suficientemente rápido como para generar electricidad de emergencia”. Necesitamos pensar sobre las posibilidades en las que la cláusula si es falsa, y no conseguimos hacerlo.

* * *

Si los seres humanos se equivocan tanto, ¿cómo pueden ser lo suficientemente racionales como para inventar la lógica y las matemáticas, y la ciencia y la tecnología? En el centro de la racionalidad humana están algunos principios sencillos que todos reconocemos: una conclusión debe ser cierta si se cumple en todas las posibilidades compatibles con las premisas. La conclusión no se sigue de las premisas si hay un contraejemplo para ella, es decir, una posibilidad que es consistente con las premisas, pero no con la conclusión. La base de la racionalidad es nuestro conocimiento de que un único contraejemplo anula una conclusión respecto a lo que tiene que ser cierto.

* * *

Gran parte de nuestro razonamiento cotidiano es difícil de evaluar a partir de un punto de partida racional. Superamos a nuestros propios programas de ordenador en la creación de explicaciones para resolver inconsistencias. Nuestra aparente fluidez puede llevarnos a pensar que estamos utilizando un procedimiento abarcable. Pero, para determinar si un procedimiento es abarcable, necesitamos saber qué computa. Necesitamos saber el rango de entradas y los resultados apropiados para ellas. No tenemos una explicación de este tipo para el razonamiento cotidiano. Por ejemplo, supongamos que usted cree:

Si a un anciano le muerde una serpiente mortal entonces se muere.

A este anciano le mordió una serpiente mortal.

Sin embargo, no murió.

Y yo le pregunto a usted: ¿cómo ocurrió? Usted podría decir:

Porque había un antídoto para el veneno.

Esta parece ser una respuesta adecuada. Supongamos, en cambio, que usted dice:

Debido a un milagro.

o:

No sé.

¿Son estas respuestas adecuadas? Ese es el problema, no tenemos una descripción de cuál se supone que tiene que ser el resultado del razonamiento cotidiano. En nuestro razonamiento sobre la disposición espacial de los objetos, sí sabemos cuál debería ser el resultado: las relaciones espaciales entre los objetos implicadas por una descripción. Pero este dominio es inabarcable. Por consiguiente, podemos enfrentarnos con los problemas a pequeña escala, pero los problemas a gran escala nos derrotan. Si el razonamiento espacial forma parte del razonamiento cotidiano, entonces el razonamiento cotidiano también es inabarcable.

Algunos filósofos y psicólogos –William James, por ejemplo– han sostenido que siempre deberíamos hacer un cambio mínimo en nuestras bellas creencias para acomodar un hecho desfavorable que entra en conflicto con ellas. Dudo que sigamos este procedimiento en todas las ocasiones, y he descrito experimentos que muestran que los participantes lo violan. Nuestra tendencia abductiva para explicar las cosas nos lleva a menudo a adoptar cambios bastante grandes en

nuestras creencias. Incluso nos lleva a pasar por alto principios estadísticos básicos, como la regresión a la media.

* * *

Unas personas razonan mejor que otras. Puede usted decir que los que razonan mejor son más inteligentes. Tiene razón en este sentido: la capacidad de razonamiento correlaciona con el rendimiento en los test de inteligencia. El problema es que nadie sabe qué miden los test de inteligencia, y muchos incluyen ítems que exigen razonamiento. Por tanto, la afirmación puede ser circular. Los psicólogos hasta ahora han aislado un factor que correlaciona con nuestra capacidad para razonar. Es la capacidad de nuestra memoria de trabajo para mantener información. Este resultado tiene sentido porque la memoria de trabajo retiene la información mientras razonamos sobre ella. Por esto es por lo que la memoria de trabajo está en el centro del poder computacional. Pero su capacidad no parece explicar todas las diferencias de capacidad entre una persona y otra. Podría ser algo tan sencillo como la práctica. Sin embargo, nadie ha descubierto todavía qué otra cosa es responsable de estas diferencias.

Ninguno de nosotros es un razonador sobrehumano. Y el cuello de botella es nuestra memoria de trabajo. Enseñarnos las leyes de la lógica no ayuda mucho a mejorar nuestro razonamiento. También se puede enseñar a los delincuentes la ley criminal para intentar mejorar su conducta. No se afronta el problema fundamental. En cambio, he descrito una forma fiable de mejorar nuestro razonamiento. El paso esencial es preguntarnos: ¿cuáles son las posibilidades compatibles con las premisas? Luego debemos tratar de enumerarlas.

* * *

Los modelos, he sostenido, subyacen en la deducción, la inducción y la abducción. Pero, ¿qué pasa con el razonamiento científico, el razonamiento matemático, el razonamiento médico, el razonamiento legal y todos los otros

tipos de razonamiento que merecen capítulos en los libros de texto sobre razonamiento? ¿Dependen de procesos mentales diferentes? La esencia de este libro es que no. La diferencia principal entre ellos está en el contenido, no en los procesos. Utilizamos la información que se nos da –las pruebas de nuestros sentidos, las descripciones y nuestro conocimiento– para imaginar qué puede estar pasando. Por supuesto, se producen procesos de razonamiento diferentes. Se nos puede pedir que extraigamos una conclusión, que evaluemos si una serie de proposiciones es o no consistente, que cuestionemos un argumento, o que determinemos si una conclusión dada es razonable a la luz de las premisas. También se producen diferencias entre la deducción, la inducción y la abducción. La inducción depende del conocimiento y las restricciones heurísticas, y la abducción depende del conocimiento para crear nuevos conceptos y explicaciones. Pero ninguna prueba que yo conozca corrobora la idea de que los científicos o los juristas adquieran procesos especiales de razonamiento alejados de los de la vida cotidiana.

Las historias de razonamiento experto que he utilizado no incluyen las humanidades, las profesiones, las artes. Ni hojalateros, ni sastres, ni soldados, ni marineros. Nadie duda de que el estudio de la filosofía, la historia, la lengua o la cultura depende de una capacidad para razonar. Los profesionales razonan, pero raramente lo mencionan. Los generales y almirantes razonan en cada batalla, pero no se refieren a ello. Las memorias de Ulises S. Grant, por ejemplo, no hablan del tema. Solo sé de una inferencia hecha por el Duque de Wellington. Antes de la batalla de Assaye en la India, que ganó aunque le superaban veinte a uno, infirió la existencia de un vado entre dos pueblos debido a que estaban construidos en lados opuestos de un río. Quizá como un legado del movimiento romántico, los artistas rara vez han comentado la necesidad de razonar para lo que ellos hacen. Como Walt Whitman, algunas veces alaban las virtudes de la inconsistencia, o como James Joyce, a veces se mantienen callados respecto a su astucia. Los autores nos dicen que escriben tumbados en la cama utilizando un portátil, o de pie en un atril con una libreta y un bolígrafo, pero no nos dicen casi nada sobre cómo piensan. Un vasto silencio respecto a su pensamiento emana de los pintores, compositores, autores teatrales, arquitectos, coreógrafos... Solo hay un puñado de excepciones. W. H. Auden, que ha sido una piedra angular en este libro, discutió el tema en varios ensayos. Stephen Spender describió cómo escribió un poema. Helen Vendler analizó los patrones de pensamiento en cuatro grandes poetas. Y Paul Valéry escribió un ensayo sobre la necesidad de razonar al componer poesía: “... si el poeta no fuera nunca más que un poeta, sin la más ligera esperanza de ser capaz de razonar de forma abstracta, no dejaría rastros

poéticos tras de sí". Mi conjetura es que la mayor parte de la creación artística se basa en la abducción, la inducción y la deducción, exactamente igual que la mayoría de los problemas que resolvemos en la vida real.

* * *

Al comienzo de este libro anuncié la fantasía de que nuestras mentes podrían ser recreadas en programas de ordenador para dudoso beneficio de la posteridad. ¿Cuánto nos falta para llegar a ese objetivo? Parte de él consiste en ser capaces de simular las formas de razonamiento. Hemos alcanzado el final del principio de la investigación que hace que las simulaciones de nuestro razonamiento sean fiables. Tenemos alguna idea sobre cómo razonamos, y conocimiento suficiente para implementar programas de ordenador que remedan ese proceso.

Es hora de llegar a una conclusión. El argumento de este libro es que nuestro razonamiento se basa en modelos mentales, y que podemos captar que una conclusión se sigue de las premisas si se cumple en todos los modelos de las premisas. ¿Es correcta esta teoría? Lo más probable es que no lo sea. Pocas teorías en ciencia, si es que hay alguna, son correctas en el sentido de decirnos la verdad y nada más que la verdad. Y ninguna teoría nos dice toda la verdad. Por tanto, podemos estar casi seguros de que la teoría es falsa en algunos aspectos. ¿Está, entonces, más cerca de la verdad que otras teorías? Quizá. Si es errónea, los psicólogos descubrirán una serie de resultados que sean contraejemplos para ella. Pero la propia teoría argumenta que los contraejemplos son fundamentales para la racionalidad. Y por tanto lo que sí puede usted decir de la teoría es: cuando se desmorone, explicará su propia desaparición.

Glosario

Abducción: Un proceso inductivo que da lugar a una explicación para un suceso o una serie de sucesos

Actitud proposicional: Una actitud mental respecto a una proposición, p. ej., creencia.

Antimodelo: Un modelo de una posibilidad en el que una proposición es falsa (ver también Contraejemplo).

Asunción por defecto: Una asunción que se puede tomar por cierta a menos que las pruebas la invaliden, p. ej., los perros tienen cuatro patas. La idea se debe a Marvin Minsky, y está incluida en ciertos lenguajes de programación de ordenadores.

Bicondicional: Una frase formada con la conectiva, "si y solo si", que es por tanto equivalente a dos condicionales, p. ej.: Si A entonces B, y si B entonces A.

Cálculo cuantificativo: El cálculo lógico que controla cuantificadores análogos a "todos" y "algunos" e incorpora el cálculo proposicional. También se conoce como "cálculo de predicados".

Cálculo de predicados: Ver Cálculo cuantificativo.

Cálculo proposicional: Una lógica de versiones idealizadas de “no”, “y” u “o”, y otras conectivas que se pueden definir en función de ellas. Esta lógica se conoce también como cálculo sentencial o lógica booleana.

Círculos de Euler: Un sistema de diagramas inventado por Leibniz (!) para representar proposiciones cuantificadas en relación con propiedades, p. ej.: “Todos los apicultores son químicos” (ver Figura 10.1), y no confundirlos con los más eficaces diagramas debidos a Venn (ver Figura 10.2).

Complejidad relacional: El número de argumentos de una relación, p. ej., Pat dio dinero a Viv tiene tres argumentos: Pat, dinero y Viv.

Condicional contrafactual: Un condicional que se interpreta como referido a una posibilidad que es contraria a la realidad, p. ej.: Si los vieneses hubieran tenido tres piernas habrían andado a paso de vals.

Conductismo: La doctrina, floreciente desde más o menos 1910 hasta 1956 en Estados Unidos, según la cual la psicología debería ocuparse de estímulos y respuestas externos y establecer leyes que los interrelacionen. La mente, en cambio, era un tema apropiado para conversaciones privadas solo entre adultos que lo aceptaran.

Conectiva: En lógica, un término que enlaza frases independientes para formar una frase nueva; en el lenguaje natural, ciertos términos como “si” y “o” desempeñan una función similar.

Conexionismo: La doctrina según la cual las teorías psicológicas deberían implementarse en modelos informáticos que simulen redes de neuronas idealizadas en el cerebro. Estas redes pueden aprender asociaciones entre estímulos y llevar a cabo inferencias simples (ver McClelland y Rumelhart, 1986). Pueden representar la información tanto en neuronas individuales como distribuida entre muchas neuronas de forma que resista daños al sistema. Un procedimiento para aprender dentro de estas redes es la propagación del error hacia atrás (desde la disparidad entre la salida real de una red y la salida requerida).

Conjunción: Una frase formada utilizando la conectiva "y" para enlazar dos oraciones principales, y por tanto con la forma gramatical: A y B.

Consistencia: Una serie de proposiciones es consistente si y solo si es posible que cada proposición de la serie sea cierta; en caso contrario, la serie de proposiciones es inconsistente.

Contradicción: En lógica, una proposición contradice a otra si y solo si una proposición debe ser cierta y la otra debe ser falsa. Es decir, una autocontradicción es una proposición como: La forma es un triángulo y no es un triángulo.

Contraejemplo: Un contraejemplo de una inferencia es una posibilidad que es consistente con las premisas pero inconsistente con la conclusión, es decir, tal como está representado en un modelo de las premisas es un antimodelo de la conclusión. Un contraejemplo de una proposición general es un ejemplo que muestra que la generalización es falsa.

Contrario: En lógica, una proposición es un contrario de otra si y solo si las dos proposiciones no pueden ser ciertas a la vez. Podrían ser las dos falsas,

p. ej.: La forma es un triángulo y es un círculo.

Cuantificador: Un término como “todos”, “algunos” o “más de la mitad”.

Determinismo: Un proceso computacional es determinista si y solo si en cada punto del proceso hay solamente una posibilidad para el paso siguiente, el cual depende de la entrada actual, si existe alguna, y del estado actual del proceso. Un proceso es no determinista si viola esta restricción, es decir, hay más de una posibilidad para el paso siguiente y no hay nada en el proceso que determine qué paso se produce. Los teóricos asumen a menudo que el sistema realiza el paso siguiente de manera correcta, es decir, “no determinismo” es un eufemismo de “mágicamente”.

Diagramas de Venn: Un diagrama eficaz para representar premisas cuantificadas (ver Figura 10.2), y no confundirlo con los círculos de Euler.

Disyunción: Una frase formada usando la conectiva “o” para enlazar dos oraciones principales. Las disyunciones incluyentes se hacen explícitas en frases con la forma gramatical: A o B, o ambos. Las disyunciones excluyentes se hacen explícitas en frases con la forma gramatical: A o bien B, pero no ambos.

Efecto atmósfera: La hipótesis de que estamos sesgados por la naturaleza de los cuantificadores de las premisas de los silogismos, p. ej.: dos premisas que contengan el cuantificador “todo” nos llevarán a una conclusión que incluya “todo”. La hipótesis fue propuesta por Sells (1936), pero tiene diversas variantes modernas.

Estrategia: Una secuencia de pasos tácticos que nos permite hacer una inferencia o resolver un problema.

Fenómeno Mary Celeste: La incapacidad para explicar un hecho.

Forma lógica: La forma lógica de una proposición representa las condiciones en las que sería cierta, es decir, su significancia. El uso de un cálculo lógico para probar una inferencia en el lenguaje cotidiano exige la recuperación de las formas lógicas de sus proposiciones. No existe ningún procedimiento general ni ningún programa informático que lleve a cabo esta tarea. Los lingüistas utilizan "forma lógica" para referirse a una representación de la parte del significado de una frase que depende únicamente del lenguaje y la estructura gramatical (ver, p. ej., Chomsky, 1995).

Frase condicional: Una frase formada utilizando la conectiva "si" para enlazar una oración subordinada A a la oración principal B, y por tanto la frase tiene la forma gramatical: Si A entonces B, o de manera equivalente: B si A.

Función de verdad: Un conectivo tiene un significado funcional de verdad si produce un valor de verdad que depende solamente de los valores de verdad de las proposiciones que conecta, p. ej., una conjunción, A y B, es cierta si y solo si A es cierto y B es cierto.

Heurístico: Un atajo o regla general que proporciona una respuesta improvisada que puede ser correcta, pero que a menudo es sesgada. Lo heurístico subyace en muchas de nuestras inferencias intuitivas.

Imágenes visuales: Una representación icónica de la apariencia de un objeto o escenario desde un determinado punto de vista.

Inabarcabilidad: Un proceso computacional es inabarcable si, a medida que el tamaño de alguna propiedad de sus entradas aumenta, el tiempo y la carga de memoria aumentan hasta un punto en que la computación deja de ser factible para cualquier dispositivo finito. Por ejemplo, valorar si una serie de proposiciones es o no consistente requiere una cantidad de tiempo y memoria imposible de manejar a medida que aumenta el número de proposiciones independientes. Las proposiciones independientes no contienen negaciones ni conectivas proposicionales, y por tanto son los constituyentes de otras proposiciones compuestas. Para encontrar una introducción general a la inabarcabilidad, véase Garey y Johnson (1979).

Inferencias ilusorias: Ver Principio de verdad.

Ingeniería inversa: La tarea de inferir cómo funciona una herramienta o un sistema a partir de un conocimiento de sus acciones observables y de cómo funcionan sus componentes.

Inteligencia artificial: La rama de la informática dedicada al desarrollo de programas que llevan a cabo tareas que requieren inteligencia.

Intuición: Un proceso heurístico inconsciente basado en premisas inconscientes que da lugar a una conclusión consciente.

Lenguaje objeto: Ver Metalenguaje.

Lógica: La ciencia de las implicaciones entre frases (a menudo en un lenguaje formalizado). Los cálculos lógicos tienen sistemas de prueba basados en reglas formales de inferencia (“teoría de pruebas”) y en una semántica (o “teoría de modelos”).

Lógica booleana: Ver Cálculo proposicional.

Memoria de trabajo: El sistema de memoria a corto plazo que mantiene la información en mente mientras trabajamos sobre ella. Esta memoria es el núcleo del poder computacional, pero la nuestra tiene una capacidad limitada y es por tanto un cuello de botella en nuestra capacidad para razonar.

Metalenguaje: Un lenguaje para hablar sobre otro lenguaje (el lenguaje objeto). En lógica, los dos lenguajes se mantienen separados. Pero un lenguaje natural, como el inglés, es su propio metalenguaje.

Metarrazonamiento: Razonamiento sobre el razonamiento.

Modelo implícito: Un modelo mental que actúa como un contenedor y por tanto sin un contenido explícito.

Modelo mental: Una representación del mundo que se propone como subyacente en el razonamiento humano; un modelo representa lo que es verdadero en una posibilidad, y hasta donde es posible tiene una estructura icónica. Los modelos mentales son el resultado final de la percepción y la

comprensión de una descripción. Los de sistemas complejos son una forma de representación del conocimiento en la memoria a largo plazo.

Modelos completos: Similares a los modelos mentales pero representando además lo que es falso en cada posibilidad.

Modulación: El uso del conocimiento sobre el significado, la referencia o el mundo para alterar la interpretación de conectivas o cuantificadores.

Negación: Un concepto que se puede expresar utilizando la palabra “no”. Una proposición negativa es cierta si y solo si la correspondiente proposición afirmativa es falsa, y viceversa.

No determinismo: Ver Determinismo.

No validez: Un inferencia es no válida si su conclusión necesita no ser cierta en caso de que sus premisas sean verdaderas. Se produce un tipo de no validez cuando una conclusión es inconsistente con las premisas, es decir, su negación es una inferencia válida a partir de ellas. Otro tipo de no validez aparece cuando una conclusión es consistente con las premisas, pero no se deduce necesariamente de ellas (ver también Validez).

Pensamiento mágico: Las inducciones de otras culturas que dan lugar a conclusiones que nosotros no creemos.

Postulado del significado: Un axioma utilizado en lógica para capturar la

propiedad lógica de un término, como la transitividad de una relación.

Potencia computacional: Lo que un sistema es capaz de computar. Un sistema más potente puede computar resultados que van más allá de los de un sistema menos potente. Dado un pequeño repertorio de procedimientos básicos entre los que se incluye la recursión, la potencia computacional depende solamente de la naturaleza de la memoria de trabajo para los resultados de los cómputos intermedios.

Presuposición: Una proposición que tiene que ser cierta para que alguna otra frase exprese una proposición, p. ej., tanto mi mujer ha dejado de fumar como su negación presuponen que mi mujer fumaba.

Principio de verdad: La asunción de que los modelos mentales representan lo que es cierto, pero no lo que es falso. Cuando la falsedad es importante, el principio proporciona inferencias falaces pero convincentes, las llamadas inferencias "ilusorias", que son una característica del uso de modelos mentales en el razonamiento.

Probabilidad condicional: La probabilidad de un suceso una vez que se produce otro, p. ej.: la probabilidad de que llueva una vez que el pronóstico es lluvia.

Proposición: Lo que una frase expresa en el contexto suponiendo que sus presuposiciones son verdaderas. Una proposición es verdadera o falsa.

Psicología evolucionista: La doctrina de que la cognición depende de "módulos" independientes del cerebro que evolucionaron hace unos 120.000 años debido a que suponían una ventaja para nuestros ancestros cazadores-recolectores. Por ejemplo, se supone que un módulo controlaría si otros nos

están engañando.

Razonamiento consciente: Un proceso iniciado de forma consciente que da lugar a una conclusión consciente, aunque el proceso en sí mismo, como todos los procesos mentales, es inconsciente.

Razonamiento deductivo: El proceso de establecer que una conclusión es una inferencia válida a partir de las premisas, es decir, debe ser cierta dado de que estas sean ciertas.

Razonamiento deóntico: Razonamiento sobre acciones que son obligatorias, permitidas o prohibidas.

Razonamiento inconsciente: Un proceso inferencial que se inicia y se lleva a cabo sin que seamos conscientes. Puede dar lugar a una conclusión consciente, una intuición, o sus consecuencias pueden ser también inconscientes.

Razonamiento inductivo: El proceso de sacar conclusiones plausibles que van más allá de la información contenida en las premisas.

Razonamiento no monotónico: La lógica es monotónica porque las premisas adicionales incrementan el número de inferencias válidas que se deducen de ellas. En la vida cotidiana, sin embargo, una premisa adicional puede llevarnos a retirar una conclusión: se dice por tanto que nuestro razonamiento es no monotónico.

Recursión: Un bucle de operaciones repetidas, como en el caso de un procedimiento que recurre a él mismo. Desempeña un papel central en computación y en el procesamiento del lenguaje. Depende de una memoria de trabajo para obtener los resultados de las computaciones intermedias.

Reducción al absurdo: Una inferencia que lleva de una suposición a una autocontradicción, y que por tanto implica la negación de la suposición (ver Capítulo 19).

Regla de inferencia específica del contenido: Una regla de inferencia formal que contiene términos que no pertenecen a la lógica sino que tienen un contenido determinado.

Regla formal de inferencia: Una regla que puede utilizarse para sacar una conclusión a partir de las premisas de manera que tiene en cuenta solamente su forma lógica, no su contenido. Los cálculos lógicos se formalizan utilizando estas reglas, y varias teorías psicológicas del razonamiento proponen que nos basamos en ellas.

Relación transitiva: Una relación que produce una inferencia de la forma A está relacionado con B, B está relacionado con C, por tanto, A está relacionado con C.

Representación icónica: Una representación en la que las relaciones entre sus partes se corresponden con las relaciones entre lo que representan, p. ej., una imagen visual.

Representación proposicional: Una representación con una estructura

sintáctica en un lenguaje mental parecido a un lenguaje natural.

Resonancia magnética funcional (RMf): Un procedimiento en el que se estudia a un participante en el experimento mientras tiene la cabeza situada dentro de un gran imán, y las observaciones dan una medida indirecta de las cantidades relativas de oxígeno consumidas en las diferentes regiones cerebrales. Aquellas regiones que son más activas requieren más oxígeno.

Semántica: Perteneciente a los significados o valores de verdad de las frases.

Significación: De una proposición, las condiciones en las que sería cierta.

Significado: El significado de una frase –excluidas las frases hechas– depende del significado de las palabras que la forman y de las relaciones gramaticales entre ellas. Los significados de las palabras en principio nos capacitan para distinguir entidades, propiedades y relaciones.

Silogismo: Una inferencia a partir de dos premisas, conteniendo ambas un único cuantificador (ver Capítulo 10).

Sistema nervioso autónomo: El sistema nervioso de todos los vertebrados que controla varios órganos, como el corazón o el estómago, en situaciones de lucha y huida, y reposo y digestión. Sus efectos son involuntarios.

Suposición: Una asunción o hipótesis hecha en beneficio de un argumento o una inferencia.

Tabla de verdad: Una tabla utilizada por los lógicos para establecer una serie de posibilidades compatibles con una o más proposiciones. También se puede utilizar para establecer el significado de una conectiva funcional de verdad: la tabla establece la verdad o falsedad de las frases formadas con la conectiva en términos de verdad o falsedad de las proposiciones que interconecta.

Táctica: Una etapa en una estrategia para hacer una inferencia o resolver un problema. Las tácticas están basadas en mecanismos inconscientes.

Tarea de selección: Una tarea de razonamiento que exige la selección de pruebas relevantes para la verdad o falsedad de una proposición –una proposición condicional en la mayoría de los experimentos (ver Capítulo 1).

Teorema de Bayes: Un teorema del cálculo de probabilidades que relaciona una probabilidad condicional con otra (ver Tabla 15.2).

Teoría normativa: Una teoría sobre cómo debemos llevar a cabo una tarea, p. ej.: el cálculo de probabilidades es una teoría normativa respecto a cómo debemos razonar sobre probabilidades.

Transiciones de Markov: Una secuencia de hechos en los que cada uno se extrae con una cierta probabilidad entre una serie finita de posibilidades. Las series de posibilidades pueden ser diferentes de una transición a la siguiente. Las probabilidades pueden depender de la secuencia previa de hechos. En principio, un proceso que genera transiciones de Markov no necesita memoria de trabajo.

Validez: Una inferencia es válida si su conclusión debe ser verdadera cuando sus premisas sean verdaderas. Un inferencia válida partiendo de premisas verdaderas da lugar a una conclusión verdadera; una inferencia válida a partir de premisas falsas puede producir una conclusión verdadera o una falsa (ver también no validez).

Valores de verdad: Una proposición tiene dos posibles valores de verdad: verdadera o falsa.

Notas en los capítulos

Los números corresponden a las páginas de la versión impresa de los capítulos a las que se refieren.

Capítulo 1: Introducción

11. Hooke estima nuestra capacidad de memoria en cien millones de ideas (Hooke, 1705, p. 144; citado por Dyson, 1997), von Neumann (1958) la estima en 1020 bits, siendo un bit la cantidad de información contenida en un dígito binario (un 1 o un 0). Landauer (1986), como Hooke, permite que olvidemos, y estima una ganancia neta a lo largo de 70 años de 109 bits. Una letra del alfabeto en un texto ocupa ocho bits, y por tanto una novela son alrededor de 107 bits. Mil novelas son 1010 bits. Los ordenadores tienen en la actualidad aproximadamente la misma capacidad de almacenamiento, p. ej., mi portátil Apple tiene 4 x 109 bits de memoria. La memoria humana trabaja de una forma diferente a la memoria de un ordenador, y es un proceso de reconstrucción a partir de fragmentos de datos.

12. La tesis Church-Turing se describe, p. ej., en Sieg (1999). Church identificaba las “computaciones” con un cálculo lógico concreto que él había ideado; Turing las identificaba con las de unas máquinas computacionales abstractas (conocidas en la actualidad como “máquinas de Turing”). Los dos sistemas resultaron ser equivalentes en el sentido de que lo que podía hacer un sistema también lo podía hacer el otro. Las tesis posteriores sobre otros sistemas también han coincidido en el mismo conjunto de computaciones. Newell (1980) sostiene que la tesis Church-Turing es la base del concepto de “procesador simbólico”, es decir, un sistema físico que puede manejar símbolos, y que los seres humanos son procesadores simbólicos. Asimismo, Hunt (1999, 2002) considera la mente como un procesador de información.

12. Descartes (1637) sostiene que el libre albedrío hace imposible una ciencia de

la mente. El argumento de Dostoyevsky aparece en Memorias del subsuelo (1864). Koestler hizo su apuesta en un periódico dominical inglés durante los años setenta.

12. El concepto de pensamiento “asociativo” se remonta a Aristóteles. El erudito victoriano Francis Galton (1883) escribió sobre sus propias asociaciones: “Quizá la más fuerte de las impresiones... considera la gran diversidad del trabajo hecho por la mente en un estado de semiconsciencia, y la razón válida que ofrecen para creer en la existencia de estratos de operaciones mentales aún más profundos, sumergidos totalmente bajo el nivel de la conciencia...”

13. Mi colega el filósofo Gilbert Harman (1973) sostiene que el pensamiento es no determinista. Hopcroft y Ullman (1979) describen el no determinismo en la teoría computacional.

13. Las definiciones de deducción e inducción en los libros de texto se remontan a Aristóteles. En Tópica (Aristóteles, 1984, Vol. 1, 105 a 13) escribe, “la inducción es un paso de particulares a universales”. Sostiene en Primeros analíticos (Libro II, párrafo 23) que cada creencia procede o de la deducción o de la inducción, y que la inducción va de lo particular a lo general. Mill (1874, p. 210) escribe: “Inducción... es esa operación de la mente, por la que inferimos que lo que sabemos que es cierto en uno o varios casos particulares, será cierto en todos los casos que se parezcan a los anteriores en ciertos aspectos determinables.”

14. El concepto de información como exclusión de posibilidades se remonta a los filósofos medievales. Bar-Hillel y Carnap (1964) resucitaron el concepto y le dieron una base precisa. Se refieren a él como “información semántica” para distinguirlo de la información en un sentido estadístico (véase Shannon y Weaver, 1949).

14. La cita de Maxim aparece en Jakab (1990, p. 26).

15. Henle (1978) escribe: “Nunca he encontrado errores que pudieran atribuirse de manera inequívoca a un razonamiento defectuoso”. Cohen (1981) sostiene que cualquier error que se produce es el resultado de fallos locales de funcionamiento en el sistema. Estas opiniones son una repetición de las de Leibniz (1765), Kant (1800) y Boole (1854).

15. Oaksford y Chater (1998) tienen con mucho la opinión más ingeniosa sobre

razonamiento y probabilidades. En el capítulo 16 discuto su descripción.

15. Revlis (1975) propone que no razonamos, sino que seguimos la “atmósfera” creada por las premisas.

15. Gabrielle Lolli, un lógico italiano, dijo en una reunión en Padua en 1999 que sin enseñarnos lógica seríamos incapaces de razonar. Hay estudios que han mostrado que esa enseñanza no nos ayuda en el caso de problemas de razonamiento nuevos (Cheng et al., 1986).

15. En cuanto a las variaciones en la capacidad de razonamiento, véase, p. ej., Johnson-Laird (1983, p. 117 et seq.).

15. En cuanto a un ejemplo sobre la capacidad de razonamiento de niños de seis años, véase Trabasso et al. (1975).

15. Las pruebas en relación con mi opinión de que los contraejemplos subyacen en nuestras inclinaciones racionales se revisan en la Parte V del libro.

15. Se cree que Aristóteles afirmó: “El hombre es el animal racional”. Lo creía, pero también que la mente contiene una parte irracional (cf. Aristóteles, Ética a Nicómaco, Libro X, Capítulo 7).

15. Stanovich (1999) muestra que la capacidad para razonar (en experimentos) correlaciona con la inteligencia (en test psicológicos).

16-17. La cita sobre el test LSAT y el ítem de la Tabla 1.1 están sacados del Official LSAT Prep Test publicado por el LSAC en 2000.

16. El testigo experto de cuya declaración tomo la cita es mi hermano Andy Johnson-Laird, un analista de sistemas forense, ocupación que él inventó. Le agradezco este ejemplo que he modificado un poco.

18. La cita de Thesiger corresponde a las páginas 65-66 de su libro Arenas de Arabia publicado por primera vez en 1959 (véase Thesiger, 1959), que describe su viaje a través del “Empty Quarter” de Arabia Saudita antes de que en ese país se descubriera petróleo.

18. Debo agradecer las inferencias de la paciente con trastorno “obsesivo-compulsivo” al psiquiatra Francesco Mancini. En el Capítulo 7 volveré a tratar

este tema.

19. La cita del LSAC aparece en el Prueba LSAT oficial.

19. La lógica de Aristóteles aparece en su Organon, una colección de textos que incluye Categorías, Tópicos, Primeros Analíticos y Analíticos Posteriores (véase Aristóteles). Los grandes pioneros de los sistemas formales de lógica son Frege (1967, publicado originalmente en 1879) y Whitehead y Russell (1927). Jeffrey (1981) es una excelente introducción a la lógica moderna, y trata tanto de la forma como del contenido. En el lenguaje simbólico de una lógica moderna, las oraciones llevan su forma lógica marcada en la frente.

19. Los lógicos consideran la forma lógica como una representación de las condiciones en las que una proposición sería verdadera, es decir, la forma lógica captura su significación (p. ej., Sainsbury, 2001, p.343). Los lingüistas, sin embargo, consideran la forma lógica como una representación del significado de las oraciones, el cual depende del lenguaje y de la estructura gramatical (véase, p. ej., Chomsky, 1995, p. 21). No existe ningún procedimiento general que permita recuperar, en ninguno de esos dos sentidos, la forma lógica a partir de oraciones en lenguaje natural.

20. Peirce (1931-1958) fue un pionero en el enfoque semántico de la lógica. Inventó las tablas de verdad que son una forma sistemática de llevar a cabo las interpretaciones de oraciones que contienen negación y conectivas lógicas, como por ejemplo “si”, “y” y “o” (véase Berry, 1952; Johnson-Laird, 2002a). Wittgenstein (1922) las reinventó. Tarski (1956) extendió el método a la lógica completa, que contiene cuantificadores como por ejemplo versiones idealizadas de “ningún” y “algunos”[1]. Hintikka (1963) y Kripke (1963) lo extienden aún más hasta la lógica modal, la que se refiere a operadores como “posiblemente” y “necesariamente”. Montague (1974) mostró cómo aplicarlo al lenguaje natural –un método adoptado por varios lingüistas y filósofos–. Una idea central es que cada regla de gramática tiene su correspondiente regla semántica. Cuando se utiliza la regla gramatical para analizar los elementos que constituyen una frase, la regla semántica se puede utilizar para “componer” los significados de estos elementos formando el significado de la frase como un todo (véase, p. ej., Partee, 1995).

20. Boolos y Jeffrey (1989) esbozan las pruebas que relacionan los cálculos formales con los sistemas de interpretación.

20. Entre los lógicos que han tenido en cuenta el contexto en sus análisis del lenguaje natural se incluyen Kamp (1981; véase también Kamp y Reyle, 1993) y Barwise (1987).

21. La cita es de Hilbert (1928/1967, p. 475). Los psicólogos Inhelder y Piaget (1958, p. 305) escriben: “El razonamiento no es más que el propio cálculo proposicional”.

21. Ruth Byrne y yo proponemos que el razonamiento deductivo humano mantiene la información y produce conclusiones parsimoniosas que exponen algo nuevo (Johnson-Laird y Byrne, 1991).

22. Sperber y sus colegas recalcan la importancia de las metas de quienes razonan (véase, p. ej., Sperber y Wilson, 1986; Sperber et al. , 1995).

22. La distinción entre lo que la mente hace y cómo lo hace se remonta a Lord Adrian, quien sostenía que si queremos entender cómo trabaja el cerebro, debemos entender primero qué está haciendo. Marr (1982) hace la misma observación en una forma que ha resultado fundamental.

23. Stanfill y Waltz (1986), Riesbeck y Schank (1989) y Kolodner (1993) exploran la teoría según la cual el razonamiento depende de un recuerdo de casos previos.

23. Boole (1854) propone un cálculo formal para la mente. Piaget publicó muchos libros y artículos, pero sus escritos son difíciles de entender. Lo esencial de su teoría según la cual el razonamiento depende de operaciones formales puede encontrarse en Inhelder y Piaget (1958) y Piaget (1965a). En el Capítulo 18 resumo su teoría sobre el desarrollo intelectual. En Inhelder y Piaget (1958), Osherson (1974-6), Johnson-Laird (1975), Rips (1983, 1994), Braine (1978), Braine et al. (1984), Cherniak (1986), Macnamara (1986), Sperber y Wilson (1986), Pollock (1989), Smith et al. (1992) y Braine y O’Brien (1998) se describen los sistemas formales de razonamiento humano. Estas teorías difieren tanto en las reglas formales de inferencia que defienden como en los procedimientos que utilizan para obtener pruebas. Fodor (1975) defiende la visión formal del “lenguaje de la mente”. Rips (1994), además, implementó su teoría en un programa de ordenador, aunque el programa ya no contiene todas las reglas. Podría utilizarse como un lenguaje de programación de aplicación general y simular, por tanto, cualquier otra teoría. En consecuencia, la teoría en

este sentido general es difícil de refutar, es decir, ningún resultado empírico podría mostrar nunca que sea falsa a menos que demostrara que los procesos mentales no son computables (Johnson-Laird, 1997). Mi preocupación en este caso no se refiere a las reglas formales que subyacen en los lenguajes de programación, sino a las teorías que postulan que razonamos utilizando reglas formales de inferencia similares a las de la lógica.

24. Wason (1966) describe sus primeros estudios sobre la tarea de selección (véase también Wason y Johnson-Laird, 1972). Wason y Shapiro (1971) describen el estudio en que muestran los efectos del contenido sobre la tarea de selección. Existe una amplia bibliografía sobre la tarea (como revisiones, véanse, p. ej., Evans et al. , 1993; Evans y Over, 2004, capítulo 5).

24. Oaksford y Chater (p. ej., 1998) sostienen que la tarea de selección no provoca razonamiento deductivo. Existen datos que muestran lo contrario (Feeney y Handley, 2000).

25. La idea de reglas específicas del contenido fue introducida en Hewitt (1969). Respecto a su utilización en los llamados “sistemas expertos”, es decir, programas para el diagnóstico médico y para otros campos, véase, p. ej., Shortliffe (1976). La regla sobre las pseudomonas aparece en Norvig (1992, pp. 531-532) y está basada en el programa de Shortliffe.

25. Searle (1980) sostiene que ningún programa de ordenador puede entender el significado de las oraciones.

26. La cita sobre los modelos mentales es de Johnson-Laird (1970a). Craik (1943) introdujo la noción en psicología, aunque propone que el razonamiento depende de reglas verbales. En Garnham (1987, 2001) y Glenberg et al. (1987) aparecen datos sobre el uso de modelos para representar el discurso. La teoría según la cual el razonamiento depende de modelos mentales se describía en Johnson-Laird (1983). Ruth Byrne y yo la ampliamos para abordar las conectivas proposicionales, tales como “si” y “o” (véase Johnson-Laird y Byrne, 1991).

26. Durante los últimos veinte años se ha producido una acumulación de datos experimentales que corroboran los modelos mentales (como referencias, véase la pagina web desarrollada por Byrne y sus colegas en el Trinity College, Universidad de Dublín: http://www.tcd.ie/Psychology/People/Ruth_

Byrne/mental_models). Los modelos han desempeñado también un papel cada vez mayor en inteligencia artificial, véanse, p. ej., Halpern y Vardi (1991) y Glasgow (1993).

26. Shastri y sus colegas describen un sistema “conexionista” para simular una red de células nerviosas idealizadas capaz de realizar inferencias sencillas (véanse, p. ej., Shastri, 1988, y Shastri y Ajjanagadde, 1993).

26-7. Las opiniones de Heidegger se describen en su libro What is called thinking? Su influencia puede encontrarse, p. ej., en la crítica de la inteligencia artificial de Dreyfus (1972). Las gallinas domésticas parecen refutar a Heidegger porque recuerdan dónde tomaron la comida y qué era (Forkman, 2000), lo mismo que hacen pollos de cinco días (Cozutti y Vallortigara, 2001).

Capítulo 2: Iconos e imágenes

29. El epígrafe pertenece al poema Palabras (Auden, 1976, p. 473)

31. Lewison (1996) revisa su trabajo con los Gugu Ymithirr.

32. Gladwin (1970) informa sobre sus experiencias aprendiendo a ser un navegante de las Islas Carolina. Oatley (1976) compara esta habilidad con otros métodos de navegación.

32. Tolman (1948) sostiene que los animales construyen mapas cognitivos de su entorno y llevó a cabo el experimento en que bloquea la ruta tortuosa hacia la caja de comida. Hull (1934) defiende las respuestas fraccionarias anticipadoras de meta.

33. Von Frisch (1966) contiene la descripción clásica de las abejas “bailarinas”. En Gould y Gould (1994) se pueden encontrar observaciones etológicas de la búsqueda de alimento por animales y abejas. Las arañas saltarinas también tienen representaciones internas de rutas que nunca han seguido (Jackson y Tarsitano, 1997).

33. Landau et al. (1984) mostraron que, después de moverse entre varios objetos,

una niña ciega infería nuevos caminos entre ellos. Su modelo del mundo incluye información métrica y geométrica dentro de un esquema de coordenadas espaciales.

33. Los datos sugieren que las regiones del hipocampo y del córtex parietal construyen las representaciones espaciales tanto egocéntricas como objetivas (véanse O'Keefe y Nadel, 1978, y Bryant, 1992 a,b). Asimismo, los pájaros que almacenan comida para consumirla más tarde tienen un hipocampo más grande que los pájaros, genéticamente relacionados con ellos, que no almacenan comida (véase Rogers, 1997), y los taxistas de Londres que tienen el "conocimiento", es decir, conocen la distribución de todas las calles de la ciudad, también tienen un hipocampo agrandado (Maguire et al. , 2000). Ciertas células del hipocampo de los animales llamadas células "de lugar" ["place" cells] solo se activan si el animal está en un lugar determinado (sin tener en cuenta los indicios referidos a ese lugar). Scoville y Milner (1957) describen al paciente conocido como "H.M.", al que se le había extirpado parte del hipocampo. Eichenbaum (1999) conjetura respecto al hecho de que el hipocampo representa relaciones entre entidades.

34. La descripción que Bernard Shaw hace del estudio de Higgins aparece al principio del Acto II de Pygmalion (véase Shaw, 1965, pp. 720-721).

34. Para encontrar datos respecto a que los modelos basados en descripciones son similares a los que se basan en la percepción, véase, p. ej., Bryant (1992 a,b). Marr (1982) sostiene que la visión produce un modelo mental que explicita qué cosas están dónde en el decorado.

35. La noción de "iconicidad" se debe a Peirce (p. ej., 1931-1958, Vol. 4, párrafo 447). Sostiene que los diagramas deberían ser tan icónicos como sea posible (Vol. 4, párrafo 433), es decir, las partes de un diagrama deberían estar interrelacionadas en la misma forma en que lo están las entidades que ese esquema representa (Vol. 4, párrafo 418). Entre otras formulaciones independientes de la iconicidad están la versión de los diagramas de Maxwell (1911), y la teoría "pictórica" del significado de Wittgenstein (1922), con su proposición clave 2.15: "El que los elementos del cuadro estén combinados entre sí de una determinada forma, representa que las cosas [en el mundo real] están combinadas del mismo modo unas con otras". Otra versión de la misma idea, cuando se produce con una forma matemática precisa, se conoce como "homeomorfismo".

36. Una biografía reciente de Leonardo da Vinci es la de Nicholl (2004).

36. La cita es de Botzmann (1974).

36. Ferguson (1992) critica a los ingenieros por su abandono de la visualización.

36. La narración del sueño de Kekulé es de Findlay (1937, p. 43). La afirmación de que la anécdota no es cierta se debe a Wotiz y Rudofsky (1984, según cita Gruber, 1994).

36. Gaston (1928) informa respecto al cuestionario sobre imagen visual. El comentario de Darwin a Galton aparece en Darwin (1887, 3.239), y es citado por Keynes (2002, p. 213).

36-7. La cita de Einstein aparece en Hadamard (1996). Wertheimer, un psicólogo, entrevistó a Einstein, y corroboró su uso de las imágenes mentales (Wertheimer, 1961, p. 228; respecto a la corrección de inexactitudes históricas, véase Miller, 1984, capítulo 5).

37. Debo agradecer los problemas de las agujas y los discos a Geoffrey Hinton.

38. Gleick (1992, p. 131) destaca el papel de las imágenes mentales en el pensamiento del difunto Richard Feynman.

38. La cita sobre la “costumbre” es de Auden (1977, p. 306).

38. Paivio (1971) postula dos modos de representación mental, uno lingüístico y el otro basado en imágenes.

39. Debo agradecer la descripción de la historia de Fisher y la degustadora de té a Salsburg (2001).

39-40. El primer estudio sobre rotación mental aparece en Shepard y Metzler (1971). Metzler y Shepard (1982, p. 45) informan sobre sus resultados respecto a que las rotaciones pueden hacerse en tres dimensiones: “Estos resultados parecen ser consistentes con la idea de que... los sujetos estaban realizando sus operaciones mentales sobre representaciones internas que eran más parecidas a objetos tridimensionales representados en dibujos bidimensionales que a los dibujos bidimensionales que realmente se les presentaban.” No todas las criaturas parecen construir ese tipo de modelos. Por ejemplo, las palomas

superan el test de rotación mental pero no muestran que la disparidad en el ángulo del eje principal del objeto entre un dibujo y otro produzca efecto sobre ellas (Hollard y Delius, 1982).

40. Hegarty (1992, 2004) informa sobre sus experimentos respecto a la simulación de aparatos mecánicos sencillos. Se producen diferencias de habilidad de un sujeto a otro (Hegarty y Sims, 1994). Los gestos que hacemos al resolver estos problemas parecen ayudarnos a seguir la pista de en qué punto del problema nos encontramos (Hegarty et al. , 2005).

41. El poder de crear imágenes mentales de Tesla parece haber sido extraordinario. Siendo niño las imágenes se interponían en su pensamiento (véase, p. ej., O'Neill, 1944).

41. El estudio sobre la revisión entre dos puntos en una imagen mental de un mapa aparece en Kosslyn et al. (1978; véase también Kosslyn, 1980).

41. Pylyshyn expresa su escepticismo sobre el papel causal de la imagen mental en varios artículos (véase, p. ej., Pylyshyn, 1973, 1981, 2003). Huxley (1893) acuñó el término "epifenómeno".

42. Entre los que proponen las representaciones enraizadas en la percepción se incluyen Barsalou (1999) y Markman y Dietrich (2000). Estos últimos escriben (p. 472): "Los argumentos teóricos y los datos experimentales sugieren que la ciencia cognitiva debería evitar las representaciones amodales."

42. Wittgenstein (1953) señala que una imagen de una gran cruz roja no nos dice en sí misma que represente una negación. Los tiempos que se necesitan para evaluar proposiciones afirmativas y negativas fueron estudiados por primera vez por Wason (1959), y más tarde por Clark y Chase (1972) y otros. Una proposición afirmativa que es cierta, como por ejemplo: "El cuadrado está a la derecha del círculo", es compatible con un tipo de posibilidad. Su falsedad es compatible con muchos tipos de posibilidad. Por el contrario, una proposición negativa que es cierta, como: "El cuadrado no está a la derecha del círculo", es compatible con muchos tipos de posibilidad. Su falsedad es compatible con un único tipo de posibilidad. Este modelo puede explicar una observación extraña: cuando verificamos proposiciones, tardamos menos en determinar que una proposición afirmativa es cierta que en determinar que es falsa, mientras que en el caso de una proposición negativa tardamos menos en determinar que es falsa

que en determinar que es cierta (Wason y Johnson-Laird, 1972).

42. La idea de que pensamos en todas las posibilidades positivas equivalentes a la negación se ha corroborado experimentalmente (Oaksford y Chater, 1994; Schroyens et al. , 2001a).

42. Peirce desarrolló dos sistemas de diagramas para la lógica, que aprovechan su estructura icónica (respecto a una descripción desde un punto de partida lógico, véase Shinn, 1994, y respecto a una descripción de sus implicaciones psicológicas, véase Johnson-Laird, 2002a). Anderson (1978) sostiene que para entender cualquier sistema de representación es necesario conocer tanto el formato de las representaciones como los procesos que operan sobre ellas.

43-4. Los estudios iniciales sobre imágenes mentales mostraron que mejoran el razonamiento sobre relaciones espaciales (p. ej., De Soto, London y Handel, 1965; Shaver et al., 1975). Los estudios que no muestran ningún efecto o muestran un efecto negativo tienden a examinar relaciones visuales (p. ej., Sternberg, 1980; Richardson, 1987; Clement y Falmagne, 1986). Egan y Grimes-Farraw (1982) observan que los participantes que informan del uso de imágenes mentales se comportaron peor que aquellos que no lo hicieron. Farah et al. (1988) informan de datos que corroboran sistemas diferentes para las imágenes y para las representaciones espaciales. Knauff et al. (2001) examinan las disrupciones de las inferencias provocadas por tareas secundarias que eran visuales o auditivas, y espaciales o no espaciales. Solo las tareas espaciales provocan una disrupción en la inferencia, independientemente de que sean auditivas o visuales.

43-5. El experimento que contrastaba las representaciones visuales y espaciales pertenece a Knauff y Johnson-Laird (2002). El estudio mediante resonancia magnética (RMf) procede de Knauff et al. (2003): el razonamiento en general activaba regiones del hemisferio izquierdo (la circunvolución temporal media), el hemisferio derecho (la corteza parietal superior) y ambos hemisferios (el precuneus y las circunvoluciones frontales inferior y media en la corteza prefrontal). Las relaciones visuales, como muestra la Figura 2.6, activan áreas en la corteza visual. Knauff y May (2006) informan sobre el estudio realizado con participantes ciegos de nacimiento.

44. La investigación sobre la manipulación de imágenes se describe en Finke y Slayton (1988).

45. El análisis del problema de Kekulé se basa en la versión de Findlay (1937) (véase p. 149).

46. La idea de que cada modelo mental representa un conjunto de posibilidades se debe a Barwise (1993).

47. La cita es de Golding (1955).

Capítulo 3: Modelos de posibilidades: de los trucos de magia a las catástrofes

49. Escribí una versión de un artículo sobre cómo encontrar objetos perdidos – inspirada en mi experiencia con el cuaderno perdido y también en la maravillosa cancioncilla de Elizabeth Bishop, One art [Un arte], con su estrofa inicial:

El arte de perder no es difícil de dominar

Tantas cosas parecen llenas de intención

de perderse que su pérdida no es una catástrofe.

Paolo Legrenzi me pidió que me reuniera con María Legrenzi y Vittorio Girotto para escribir un artículo en memoria de la psicóloga italiana Mimma Peron (véase Legrenzi et al. , 2001-2002). En Klein (1998, pp. 267-268) se presenta una opinión similar sobre cómo encontrar objetos perdidos.

50. Los filósofos sostienen opiniones diferentes sobre actitudes proposicionales tales como fe e incredulidad. Algunos las descartan como mitos de “psicología popular” (p. ej., Stich, 1983), mientras otros las tratan como mapas mediante los que nos guiamos (p. ej., Fodor, 1981; Ramsey, 1990).

50. Las etiquetas en los modelos son similares a la idea de “evaluaciones” de

Jackendoff (1996), que él trata como propiedades, como por ejemplo “conocido” frente a “nuevo”.

50. Véase la nota al Capítulo 1 para encontrar referencias en relación con las teorías del razonamiento basadas en reglas formales de inferencia.

50. En cuanto a una historia de las lógicas “modales”, que abordan las posibilidades, véase, p. ej., Kneale y Kneale (1962), y si se quiere una introducción a los sistemas modernos, véase, p. ej., Hughes y Cresswell (1996). Mi colega Dan Osherson fue uno de los primeros psicólogos que desarrolló una teoría del razonamiento sobre posibilidades (véase Osherson, 1974-76).

51. La literatura psicológica sobre la memoria de trabajo es amplia. Una descripción accesible debida a uno de sus principales investigadores es Badeley (1986). En cuanto a la relación entre memoria de trabajo y potencia computacional, véase, p. ej., Johnson-Laird (1993). Simon (1983) señala que la memoria de trabajo es el cuello de botella en el razonamiento.

51. Hopcroft y Ullman (1979) explican por qué la potencia computacional depende de una memoria para los resultados de las computaciones intermedias.

51. Johnson-Laird et al. (1992) y Bauer y Johnson-Laird (1993) informan de los experimentos sobre razonamiento a partir de parejas de disyunciones.

51-2. Las cinco posibilidades compatibles con las premisas:

Jules está en Gales o Charles está en Escocia, o ambas cosas

Charles está en Escocia o Kate está en Irlanda, o ambas cosas

son:

June está en Gales Charles está en Escocia Kate está en Irlanda

June está en Gales Charles está en Escocia

June está en Gales Kate está en Irlanda

Charles está en Escocia Kate está en Irlanda

Charles está en Escocia

52-3. Rips (1994, pp. 365-369) informa de su estudio sobre inferencias conjuntivas y disyuntivas.

54. García-Madruga et al. (2001) informan de los experimentos en los que establecen, por ejemplo, que las inferencias conjuntivas son más fáciles que las disyuntivas a menos que los participantes puedan trabajar partiendo de las conclusiones.

54. La teoría de modelos postula que podemos hacer suposiciones, pero investigamos sus consecuencias en modelos mentales (Van der Henst et al. , 2002). Victoria Bell llevó a cabo un estudio no publicado en el que los participantes respondieron a preguntas sobre las consecuencias posibles y necesarias de diversos hechos hipotéticos.

55-6. El estudio respecto el razonamiento sobre los partidos de baloncesto uno contra uno es de Bell y Johnson-Laird (1998). A diferencia de los adultos, los niños no son reacios a decir que un suceso podría producirse aunque sea obvio que debe producirse (Noveck, 2001). Los niños se parecen más a los lógicos, porque son menos sofisticados en cuanto al uso informativo del lenguaje.

56. Neisser y Weene (1962), p. ej., informan de la dificultad que supone aprender conceptos basados en disyunciones excluyentes. En el Capítulo 24 se describe la dificultad que supone crear diagramas de cables basados en ese tipo de disyunciones.

56. Wason y Johnson-Laird (1972) presentan un estudio que muestra la dificultad para entender documentos como el “Subsidio de defunción”. La señora Tatcher abolió el subsidio, pero en el momento de escribir esto todavía existe en Jersey, una de las Islas del Canal.

56. Los estudios que muestran que las alternativas disyuntivas provocan dificultades en la toma de decisiones fueron llevados a cabo por mi colega Eldar Shafir y el difunto Amos Tversky (Shafir y Tversky, 1992; Tversky y Shafir, 1992). Los psicólogos saben desde hace tiempo que las disyunciones provocan dificultades al razonar y elegir (Newstead y Gringgs, 1983).

57. El truco Mascot Moth se describe en Steinmeyer (2003, pp. 191-192).

57. Reason (1990) describe los factores psicológicos de la catástrofe en la estación nuclear de Three Mile Island.

Capítulo 4: La arquitectura mental y el inconsciente

59. El epígrafe de la Parte II pertenece al poema El laberinto (Auden, 1976, p. 237).

61. Todorov et al. (2005) informan del estudio sobre los juicios respecto a la competencia a partir de fotografías y su relación con los resultados de las elecciones.

62. Freud (1973) postula procesos del pensamiento irracionales (primarios) y racionales (secundarios).

62. Helmholz (1962) se refiere a las inferencias inconscientes en la percepción (véase también Whyte, 1978, para encontrar una historia del inconsciente anterior a Freud).

62. Pascal en su Pensamientos (1966, p. 211) escribe: “Es raro que los geómetras sean sutiles y los sutiles sean geómetras, porque los geómetras quieren tratar geométricamente los temas sutiles, y se vuelven ridículos, al querer comenzar por las definiciones y seguir por los principios, manera de obrar que no es la adecuada para esta clase de razonamiento. No es que el espíritu no lo haga, sino que lo hace tácita, natural y simplemente, pues la expresión de esto supera a todos los hombres y el sentimiento no pertenece más que a unos pocos.”

63. Entre los psicólogos que defienden sistemas duales de razonamiento se

incluyen Reitman (1965), Wason y Evans (1975), Johnson-Laird y Wason (1976, pp. 5-6), Johnson-Laird (1983, pp. 127 et seq.), Sloman (1996), Stanovich (1999), Rader y Sloutsky (2002) y Kahneman y Frederick (2002). Evans y sus colegas, no obstante, lo han convertido en su finca particular, véanse, p. ej., Evans (1984, 1989, 2003), Evans y Over (1996). Nisbett (p. ej., 2003) distingue también entre razonamiento "holístico", que él supone es prerrogativa de los orientales, y razonamiento "analítico", que él supone es prerrogativo de los europeos. La gran dificultad es precisar el mecanismo para el sistema rápido, intuitivo e inconsciente. Evans y Over (1996), por ejemplo, distinguen entre racionalidad1, que alcanza nuestras metas cotidianas, y racionalidad2, que alcanzamos cuando tenemos razones para nuestros pensamientos que están sancionadas por una teoría normativa. Pero no especifican qué computa el sistema para la racionalidad1 o cómo lleva a cabo estas computaciones. Comentan, sin embargo, que es: generalmente fiable y eficiente para conseguir las propias metas, extremadamente potente desde el punto de vista computacional, rápido, preconsciente, de naturaleza conexionista, no todo o nada puesto que da lugar a la racionalidad1 como una propiedad gradual, casi automático, modelado por la experiencia pasada y fuera del control de la consciencia, basado en procesos tácitos y paralelos, y compartido con los animales que pueden ser de tipo racional1 (pp. 8, 10, 22, 35, 36, 96, 143 y 154). Rips (2002, p. 408) comenta: "No es juego limpio decir que algo es una teoría de dos sistemas, por ejemplo, cuando en realidad es una teoría bien definida [para el razonamiento deliberado] más todo lo demás (Bedford, 1997).

63. Rips (2002, p. 396) señala el problema que supone obtener pruebas que corroboren la existencia de dos sistemas de razonamiento frente a la existencia de un solo sistema que actúa con dos niveles de eficiencia distintos (véase también Sloman, 1996). La teoría de modelos proporciona dos niveles de ese tipo (véase Schroyens et al. , 2003, en relación con esta idea). En el Capítulo 5, propongo un mecanismo para las inferencias inconscientes; a diferencia de la propuesta de Evans y Over es computacionalmente más débil que el sistema para las inferencias conscientes, pero puede hacer frente a información en paralelo.

63. La idea de que no somos conscientes de ningún proceso mental, sino solamente de sus resultados, se remonta a Lashley (1958). Jackendoff (1996) hace la misma observación.

64. Johnson-Laird (1983) describe la arquitectura jerárquica, postula también modelos mentales del yo y explica su papel en la conducta deliberada. Oatley

(1978) incluye una destacada revisión de la historia de las ideas jerárquicas sobre la organización del cerebro. Mi antigua colega Marcia Johnson (1983) propone una arquitectura mental que también incluye múltiples procesos en paralelo. Ella y sus colegas sostienen que los modelos con un único ejecutivo no pueden explicar adecuada y suficientemente diversos aspectos de la consciencia, tales como consciencia y control, y las formas en que falla la consciencia (véase, p. ej., Johnson y Reeder, 1997). Esta opinión puede ser correcta, pero no insistiré en ella aquí. En lo que estamos de acuerdo es en que los procesos inconscientes son computacionalmente menos potentes que aquellos que tienen acceso a la consciencia. En especial, el modelo de Johnson permite el 'pensamiento recursivo sobre el propio pensamiento' y 'la reflexión sobre las propias reflexiones', desde el punto de vista de dos sistemas que interactúan.

65. Hopcroft y Ullman (1979) muestran que las computaciones en serie y en paralelo no difieren en potencia, es decir, en cuanto a lo que pueden computar. Estos autores describen también los problemas inabarcables en computación (véase también Poundstone, 1988).

65. Malthus publicó su Ensayo sobre el principio de población en 1798 (véase Malthus, 2004).

67. En cuanto a una descripción de la psicología de la memoria de trabajo, véase, p. ej., Baddeley (1986). Baddeley (2000) añadió en su teoría un nuevo componente a la memoria de trabajo para explicar ciertos fenómenos de la consciencia.

67. Minsky (1968) sugiere que un mecanismo computacional que se autodescribe podría utilizar su propia descripción para predecir lo que haría en circunstancias hipotéticas.

67. El psicoanalista John Bowlby destaca el papel de los modelos mentales que tenemos de nosotros mismos a medida que nos desarrollamos (véase, p. ej., Bowlby, 1988).

68. La máxima autoridad en expresiones faciales, Paul Ekman, distingue entre sonrisas voluntarias e involuntarias (véanse, p. ej., Ekman, 1980, 1992b). La sonrisa espontánea depende de la contracción de los músculos que elevan las comisuras de los labios y fruncen los ojos.

68. El efecto "cóctel" está descrito en Cherry (1980).

68. En el Capítulo 6 discuto nuestra incapacidad para controlar nuestras emociones.

68. El comentario de Paul Valéry sobre el conocimiento aparece en Analects (véase Valéry, 1970, p. 405).

Capítulo 5: Intuiciones y razonamiento inconsciente

69. El epígrafe pertenece al poema Carta de Año nuevo (Auden, 1976, p. 169).

70. Nisbett y Wilson (1977) sostienen que muchas introspecciones son racionalizaciones. Los estudiosos del razonamiento, sin embargo, han encontrado que cuando los participantes en un experimento piensan en voz alta a la vez que razonan, sus descripciones parecen ser correctas (Ericsson y Simon, 1980; Van der Henst, Yang y Johnson-Laird, 2002).

70. Muchos psicólogos proponen descripciones del razonamiento de "proceso dual", contraponiendo el razonamiento intuitivo con el razonamiento deliberado (véanse las notas al capítulo anterior). La presente explicación distingue cuatro tipos de razonamiento, porque incluye las inferencias totalmente inconscientes.

70. Ekman (p. ej., 1995) ha estudiado las emociones fugaces que se entrometen durante un instante cuando tratamos de suprimir nuestros verdaderos sentimientos.

70-1. El estudio en que las palabras tienen un efecto de priming sobre conductas inconscientes, como caminar más lentamente, aparece en Bargh et al. (1996).

71. Steele y Aronson (1995) informan del experimento en el que la identificación de la propia pertenencia a una raza puede afectar a la actuación en el GRE. Shih et al. (1999) muestran que el priming fue inconsciente. A un grupo de mujeres asiático-americanas participantes en su estudio se le planteó una cuestión, como: ¿Qué idiomas habla usted en casa? A otro grupo se le preguntó: ¿Prefiere una casa mixta o un piso solo para un solo sexo? En la primera pregunta el priming afectaba al reconocimiento de la raza de las participantes, mientras que en la segunda el priming afectaba a su género. Los asiático-

americanos tienen fama de ser buenos en matemáticas; pero las mujeres tienen fama de ser malas en matemáticas. Los correspondientes efectos de priming se ponían de manifiesto en el comportamiento posterior de los grupos al realizar un test matemático.

71-2. El estudio de las asociaciones, por ejemplo, entre “mujer” y “emprendedora”, forma parte del Test de Asociación Implícita (véase Greenwald et al. , 2003). Mide las asociaciones inconscientes. Puede hacer usted mismo el test en http://moral.wjh.harvard.edu. En cuanto a valorarlo escépticamente como medida de prejuicios implícitos, véase Blanton y Jaccard (2006).

72. Muchos estudios han examinado cómo controlan los participantes los sistemas complejos, como por ejemplo simulaciones de una ciudad o una fábrica (véase, p. ej., Broadbent et al. , 1986) Pueden aprender a controlarlos sin un conocimiento consciente de cómo funciona el sistema, pero tienden a equivocarse con sistemas que dependen de múltiples variables (véase Dörner, 1996).

72. La anécdota sobre Peirce es de Sebeok y Umiker-Sebeok (1983). Fiedler y Schenck (2001) y Uleman et al. (2005), p. ej., presentan estudios sobre impresiones inmediatas de los participantes a partir de breves descripciones de cuadros.

72. Klein (1998) describe las inferencias inmediatas de los jefes de bomberos.

72-3. Hull (1920) describe el estudio sobre los ideogramas chinos.

73. Maier (1931) informa del estudio en el que los participantes tenían que atar dos cuerdas que colgaban del techo. Según Schooler et al. (1993), el pensar en voz alta empeora el insight.

73. Las inferencias sobre las referencias de los pronombres se discuten en Garnham (2001).

74. La memoria de trabajo se discute en el Capítulo 3.

74. Debo agradecer el ejemplo de The Tin Men de Michael Frayn a George A. Miller, que lo utiliza para describir un tipo de gramática que él investigó en su proyecto “Grammarama” (véase Miller, 1968, capítulo 7). Estas gramáticas llamadas “regulares” pueden se características del resultado de una máquina de

estado finito de Markov.

75. Hopcroft y Ullman (1979) describen las “máquinas de estado finito de Markov con límite k”.

76. Levesque (1986) discute oraciones basadas en la conjunción y en afirmaciones universales. Se refiere a ellas como oraciones “vívidas”, porque a partir de ellas podemos formarnos una única representación.

76. El capítulo 22 describe cómo construimos cadenas causales a partir de los eslabones elementales de conocimiento causal de la memoria a largo plazo.

76. El estudio sobre las palomas que atribuyen cuadros se debe a Watanabe et al. (1995). Ganó el premio Ig Nobel de Psicología en 1995. Los premios Ig Nobel se dan a estudios que hacen reír a la gente, y luego hacen pensar.

76-7. Marcus et al. (1999) informan sobre el estudio con bebés que aprenden lenguajes artificiales sencillos. Marcus (2001, capítulo 3) describe la dificultad para explicar estos resultados utilizando teorías del aprendizaje “conexionistas” y ortodoxas que representan solo sílabas determinadas.

77. Las inferencias basadas en premisas, como por ejemplo, “Todos tenemos prejuicios contra una persona con prejuicios”, se describen en el Capítulo 11.

77. En cuanto a un estudio sobre la publicidad subliminal, véase, p. ej., Kihlstrom et al. (1992). De acuerdo con William James (1890, p. 526), “Podemos establecer como cierto que toda representación de un movimiento despierta en algún grado el movimiento real al que se refiere”.

77. Michotte (1963) descubrió las inferencias automáticas respecto a que un objeto causa que otro se mueva.

78. Véase el Capítulo 13 en relación con la inabarcabilidad de la inducción.

78. La descripción clásica del pensamiento mágico aparece en Fraser (1890). Scoditti (1984) estudió a los talladores de madera en Kitawa. Evans-Pritchard (1937) informa sobre sus famosos estudios respecto a los Azande. Bowker (2003) cuenta la historia de Orwell y la magia.

79. Respecto a los ejemplos de pensamiento mágico en la medicina occidental,

véase Malleson (1973) y, aunque él no presta atención a este aspecto del pensamiento de los médicos, la extraordinaria historia de la medicina de Porter (1997).

79. Sutherland (1976) cuenta el caso de la máquina de TEC estropeada.

79. Gigerenzer y sus colegas describen los heurísticos rápidos y frugales (véase, p. ej., Gigerenzer y Hug, 1992; Gigerenzer et al., 1999; Gigerenzer y Selten, 2002).

80. Timothy Wilson y sus colegas han llevado a cabo diversos estudios en los que la satisfacción de los sujetos respecto a una opción era menor si habían pensado en ella que si no lo habían hecho, p. ej., la elección de pósteres de arte. Escribe: “El truco está en reunir suficiente información como para desarrollar una reacción visceral informada y luego no analizar demasiado la sensación” (Wilson, 2002, p. 172). Dijksterhuis et al. (2006) informan de un estudio similar que examinaba las hipotéticas elecciones entre coches descritos a partir de cuatro o doce características. Gladwell (2005) describe juicios realizados en un abrir y cerrar de ojos. Tall (2006) es una parodia: “Ya sabe usted de qué trata este libro. El poder de BLANK: el poder de no pensar en nada en absoluto. Utilizando lo que los investigadores científicos llaman (o llamarían si realmente les importara investigarlo) ‘Cortar láminas extrafinas’, tu cerebro ya ha decidido si te va a gustar BLANK, si su cubierta combina con tu camisa, y si te hará parecer inteligente si alguien te ve leyéndolo en el tren.” No confesaré, como me dijo mi hermano, que no corrí a no comprar un ejemplar del libro de modo que no pudiera leerlo y beneficiarme de no pensar en él.

80. Los problemas de dirigir sistemas complejos, como por ejemplo simulaciones económicas, sin pensar lo suficiente se describen en Dörner (1996). Alude a la “lógica del fallo” en nuestra tendencia a centrarnos en las cosas de una en una, por ejemplo, y a nuestra incapacidad para captar las dependencias. Vuelvo a tratar este último problema en el Capítulo 24 referido a cómo resolvemos problemas.

Capítulo 6: Las emociones como inferencias

81. El epígrafe pertenece al poema Lingüística natural (Auden, 1976, p. 636).

81. Simon (1982) recalca la naturaleza “limitada” de la racionalidad humana.

82. El capítulo 4 discute la inabarcabilidad del razonamiento (véase también el Glosario). Neisser (1963) señaló los problemas que surgen de los objetivos múltiples.

82. El comentario de Aristóteles sobre los insultos que nos hacen enfadar aparece en Política (1984, línea 1312b30, en la espléndida edición de Barnes). Aristóteles es el bisabuelo de la teoría cognitiva de las emociones (véase su versión de las emociones en Retórica, Libro II, y Ética a Nicómaco, p. ej., línea 1106a3). Reconocía que las emociones tienen una expresión facial característica (Physiognomics, línea 805b1). También sostenía que las emociones pueden distorsionar nuestras percepciones (On Dreams, línea 4601b) y nuestros juicios (Retórica, Libro I, línea 1356a14). Véanse Fortenbaugh (1975) y Elster (1999) para encontrar versiones modernas del análisis cognitivo de las emociones de Aristóteles.

82. William James (1884) describe su teoría de la emoción. Bermond et al. (1991) muestran que la parálisis puede mejorar las emociones.

82. En relación con la función potencial de las emociones véanse, p. ej., Minsky (1985) y Nesse (1991). Platón en Timeo (69d) ofrece una letanía de quejas en contra de las emociones. Gurdjieff (1991), un místico ruso del siglo veinte, sostiene que la emociones negativas, como la angustia y el miedo, consumen energía que en caso contrario podría ser dedicada a emociones positivas. En Ekman et al. (2005) aparece una descripción de las opiniones psicológicas de los occidentales y los budistas tibetanos sobre las emociones. Los estoicos, tanto los fundadores griegos del movimiento como sus seguidores romanos, premiaban la virtud por encima de la emoción. Debemos ser amables con nuestros hijos, no porque les queramos, sino porque eso es lo que hay que hacer (véase Murray, 1915). Sartre (1938) sostiene que las emociones son pensamiento ilusorio. Pocos de los autores citados han tenido algo que decir respecto a las emociones en los demás mamíferos sociales.

83. Ekman (1980, 1992b) distingue entre sonrisas espontáneas y esfuerzos voluntarios para reproducirlas.

83. La cita es de Darwin (1873).

83. La teoría comunicativa de las emociones se describe en Oatley y Johnson-Laird (1987), y en Oatley y Johnson-Laird (1996) aparece una versión revisada. Las relaciones entre emociones innatas básicas y emociones que dependen de las normas sociales de una cultura se revisan en Johnson-Laird y Oatley (2000).

83-4. Las expresiones faciales de emoción parecen ser universales, pero con algunas modificaciones culturales locales, véase Ekman (p. ej., 1980) y Marsh et al. (2003).

85. Muchos teóricos postulan la existencia de emociones básicas, p. ej., Ekman (1992a), Stein et al. (1993) y Power y Dalgleish (1997). Pero otros las rechazan, sosteniendo que las emociones están formadas por componentes que no se pueden reducir a emociones, como por ejemplo las respuestas autónomas y las evaluaciones cognitivas (Ortony et al., 1988; Ortony y Turner, 1990), los prototipos (Russell, 1991) o las tendencias hacia ciertas acciones (Fridja, 1986). Schacter y Singer (1962) proponen que la excitación general es común a todas las emociones. La tesis opuesta, que defendemos Oatley y yo, es que cada emoción básica tiene su propio patrón de actividad nerviosa, de neurotransmisores y de respuestas hormonales (véase también Ekman, 1992a).

86. Freud (1915) sostiene que las sensaciones emocionales nunca pueden ser inconscientes.

86. Tranel (1997) describe la amígdala. LeDoux (1996) resume las pruebas sobre el papel de la amígdala en las emociones básicas. Calder et al. (1996) informan de la afectación en el reconocimiento de expresiones faciales. Buchanan y Adolphs (2003) revisan los datos que se refieren a la amígdala y la memoria para sucesos emocionales.

87. LeDoux (1996) describe las dos vías neurales hacia la amígdala que subyacen en el miedo.

87. Sifneos (1996) describe la alexitimia. Se evalúa utilizando diversos cuestionarios de autoinforme, incluida la escala de alexitimia de Toronto. Los estudios que utilizan la resonancia magnética funcional (RMf) han detectado un reducción significativa en la actividad de una región del cerebro, el córtex posterior cingulado, que está implicada en la memoria y la emoción (véase Aleman, 2005).

87. Rozin (1996) describe su trabajo sobre el asco.

87. Darwin (1839) cuenta su aventura con la carne enlatada (véase la entrada de su diario correspondiente al 19 de enero de 1833).

88. Bowlby (1988) describe el papel de los modelos del yo en el normal desarrollo de los seres humanos.

88. Knight y Grabowecky (1995) subrayan el papel del córtex prefrontal al enfrentarse a posibilidades "contrafactuales" respecto a lo que ha sucedido.

88. Phineas Gage tiene su propia página web: www.deakin.edu.au/hbs/GAGEPAGE/ mantenida por el mayor experto mundial en su caso, Malcolm Macmillan.

89. Damasio (1994) describe al paciente moderno similar a Phineas Gage, y discute el papel de la amígdala en las emociones básicas, a las que se refiere como emociones "primarias". Sobreviven al daño en la región "ventral media" de los lóbulos prefrontales del cerebro, mientras que las emociones complejas no lo hacen. Damasio se refiere a las emociones complejas como "secundarias", y resume las relaciones entre los dos tipos de emociones de una forma que corrobora lo que sostiene la teoría comunicativa. Escribe: "Las representaciones disposicionales prefrontales, adquiridas, que son necesarias para las emociones secundarias, son un conjunto distinto de las representaciones disposicionales innatas que precisan las emociones primarias. Pero... las primeras necesitan a las ultimas para expresarse" (Damasio, 1994, p. 134).

89. El estudio sobre las decisiones de los pacientes con daño prefrontal en el juego experimental de apuestas aparece en Bechara et al. (1994).

90. Aristóteles escribe sobre nuestra incapacidad para elegir qué emoción sentir en su Ética a Nicómaco (Aristóteles, 1984, línea 1106a3).

90. Los comentarios de San Agustín sobre las erecciones aparecen en su Ciudad de Dios (véase, p. ej., 13.13, 14.16 de Agustín, 1985). La desobediencia del pene, escribe, es un castigo adecuado para la desobediencia de Adán. (¿Qué pasa con Eva?)

90. Algunos psicólogos sostienen que podemos controlar nuestras emociones (véase, p. ej., Ochsner y Gross, 2005). A veces podemos desviar nuestra atención de sucesos emocionales y a veces podemos reevaluarlos para modular nuestras sensaciones. Una brillante demostración de este tipo de efecto es un estudio que

combina el Test de Asociación Implícita, que describí en el capítulo anterior, con la imagen cerebral (Cunningham et al., 2004). Cuando los participantes blancos examinaban imágenes de rostros negros frente a imágenes de rostros blancos, su amígdala estaba más activa. El efecto era mayor cuando los rostros se presentaban durante un intervalo tan breve que los participantes decían que solo habían visto la imagen que aparecía después de un rostro. Era también mayor para aquellos sujetos que mostraban un sesgo mayor en el Test de Asociación Implícita, es decir, asociaban más a los "negros" que a los "blancos" con palabras negativas. Cuando los rostros se presentaban durante más de medio segundo de manera que los participantes podían ver qué estaban examinando, tomaban el control las regiones del córtex frontal, que median en nuestra capacidad de autocontrol. Eran más activas cuando los participantes examinaban los rostros negros y las respuestas de la amígdala disminuían. El estudio muestra que una emoción básica puede ser desencadenada por una evaluación cognitiva muy rudimentaria –de la que no somos conscientes–, pero la emoción subsiguiente se puede modular mediante una evaluación consciente. Asimismo, tenemos algún control voluntario sobre la expresión de nuestras emociones. Podemos ser capaces de suprimirlas o aumentarlas utilizando diversas estrategias cognitivas, como la distracción o la reevaluación, pero parece que no tenemos control sobre la aparición de una emoción o sobre qué emoción básica sentimos. Y no podemos desconectar una emoción como hacemos con una bombilla.

90. En relación con el "filtrado" emocional, véase Ekman (1995).

91. Shackespeare tiene el comentario de Hamlet sobre las emociones reales respecto a hechos irreales en el soliloquio sobre el actor: "¿Qué es Hecuba para él, o él para Hecuba, que le hace llorar?" (Hamlet, Acto II, Escena 2). Dadlez (1997) utiliza este comentario como título de su libro filosófico sobre el tema.

91. Los psicólogos no se ponen de acuerdo sobre el papel de la inferencia en la creación de emociones. Zajonc (1980) sostenía que las preferencias no dependen de inferencias, pero Lazarus (1982) replicaba que no dependen de inferencias conscientes.

91. En relación con los estudios que muestran que los participantes se equivocan sobre las sensaciones que anticipan, véase, p. ej., Schkade y Kahneman (1998), Gilbert et al. (1998) y Gilbert y Wilson (2000).

91. Kahneman y Tversky (1979) describen su teoría de las "perspectivas". Entre

los que también aceptan los efectos de las emociones sobre las decisiones se incluyen Frank (1988) y Mellers et al. (1988). Bar-Hillel y Neter (1996) y Mellers y McGraw (2001) informan sobre las pruebas respecto al papel que tienen emociones anticipadas sobre la elección.

92. Debo agradecer el ejemplo sobre la tosferina a Baron (2000).

92. Reason (1988) describe el papel del estrés en los errores cognitivos.

92. Oaksford et al (1996) demuestran que ver películas emotivas produce efectos preocupantes en la tarea de selección. Blanchette y Richards (2004) muestran que las palabras emotivas incluidas en premisas condicionales provocan un efecto preocupante similar sobre el razonamiento.

93. Duncker (1945) describe el problema de la vela y la “fijeza funcional”. Isen et al. (1987) informan sobre el efecto de las películas emotivas en su solución. Muchos estudios de Isen y sus colegas y de otros investigadores muestran que la felicidad mejora la creatividad (véase también, p. ej., Greene y Noice, 1988).

93. Johnson y Tversky (1983) demuestran los efectos de la ansiedad sobre la evaluación del riesgo por parte de los participantes. En relación con estudios sobre los efectos de las emociones sobre la vigilancia, véanse Broadbent y Broadbent (1988) y Mathews et al. (1997), y en relación con sus efectos sobre la memoria, véase Teasdale (1988).

93. Alloy y Abramson (1979, 1988) comparan las estimaciones del control sobre los resultados de participantes ligeramente deprimidos con las de participantes que no están deprimidos.

93. Power y Wykes (1996) demuestran los efectos de la depresión sobre el razonamiento respecto a asuntos positivos y negativos.

94. El estudio sobre la historia de Russell Banks fue realizado por Oatley, Nundy y Larocque (véase Oatley, 1996). Nundy (1995) mostró también que los participantes contentos eran más exactos en el razonamiento con silogismos que los participantes tristes. Indujo las emociones utilizando fragmentos de películas.

94. Aristóteles (1984) escribió sobre el enojo en Retórica (Libro II, línea 1378a).

94. Kahneman y Tversky (1982) discuten los pensamientos contrafactuales que

uno tiene cuando los planes no funcionan y las diferencias al imaginar diferentes tipos de posibilidad contrafactual (véase también Kahneman, 1995). Girotto et al. (1991) muestran que los participantes tienden a deshacer mentalmente las acciones deliberadas del protagonista.

94. El libro de Ruth Byrne (2005) The Rational Imagination describe el trabajo de la imaginación en la creación de posibilidades contrafactuales (véase también Byrne, 1996, 1997, 2002). Deshacemos sucesos excepcionales o aquellos que están bajo el control del protagonista (Kahneman y Miller, 1986; Gavanski y Wells, 1989; Girotto et al., 1991; Klauer et al., 1995).

94. Respecto a las relaciones entre emociones complejas, como el arrepentimiento y la culpa, y los pensamientos sobre posibilidades contrafactuales véanse, p. ej., Kahneman y Tversky, 1982; Niedenthal, Tangney y Gavanski, 1994; Byrne y McEleney, 2000; Byrne, 2005. Medvec et al. (1995) muestran que ese tipo de pensamientos pueden mejorar los resultados, por ejemplo, si solo has ganado una medalla de bronce en las Olimpiadas. Roese (1997) informa de un estudio en el que se completaban expresiones contrafactuales después de ver un fragmento triste de Bambi. Los que veían el fragmento completaban más rápidamente que los participantes de control una expresión referida a su actuación en una prueba previa relacionada con su habilidad para resolver anagramas que comenzaba: “ojalá...”

95. Darwin (1873, p. 38) describe su aventura con la serpiente en el Zoo de Londres.

Capítulo 7: El razonamiento en las enfermedades psicológicas

97. Las dos descripciones proceden de un estudio llevado a cabo por el psiquiatra italiano Francesco Mancini y su colega Amelia Gangemi en la que 34 psiquiatras tuvieron que diagnosticar a pacientes a partir de estas y otras descripciones. Los psiquiatras no tuvieron dificultades para identificar la naturaleza de la enfermedad psicológica en ambas descripciones. Mancini, Gangemi y yo colaboramos en los estudios de razonamiento y en el desarrollo de la teoría de las enfermedades psicológicas que se describe en este capítulo (véase Johnson-Laird et al., 2006b).

98. Szasz (1974) es escéptico respecto a la existencia de enfermedades mentales.

98. El concepto original de “neurosis” era el de una enfermedad mental que tenía una causa psicológica más que neurológica, y que, a diferencia de, pongamos por caso, la esquizofrenia, no conducía a una total falta de contacto con la realidad. La Asociación Americana de Psiquiatría, que ya no reconoce una categoría de “neurosis”, publica el Diagnostic and Statistical Manual of Mental Disorders DSM-IV-TR (2000) [Manual Estadístico y Diagnóstico de Transtornos Mentales]. Los transtornos equivalentes a las neurosis se clasifican en muchas categorías, incluidos los Transtornos de Humor y los Transtornos de Ansiedad.

98. Descartes, el arquetipo dualista, sostiene que la glándula pineal conecta la mente con el cuerpo (véase Descartes, 1984-91).

98. La investigación de John Snow sobre la transmisión del cólera se describe en el penúltimo capítulo del libro.

99. Herink (1980) describe 250 tipos diferentes de psicoterapia.

99. En la reducción de la ansiedad parecen tener éxito diversos tipos de psicoterapia y terapia de conducta (Smith et al., 1980). Consumer reports [Informes del consumidor], el periódico que informa sobre los productos manufacturados, da un valor positivo de la psicoterapia (en su número de noviembre de 1995). Este informe a su vez tiene una evaluación positiva por parte de Seligman (1995). La terapia cognitiva para la depresión, tal como se describe en Beck et al. (1979), se considera mejor que los medicamentos antidepresivos (Robinson et al., 1990). Hollon et al. (2002) revisan también diversas formas de terapia para la depresión e informan de que la terapia cognitiva de la conducta es más eficaz que una versión menos intensa del psicoanálisis freudiano, y que refuerza la terapia con medicamentos.

99. La terapia de conducta se describe en Wolpe (1990). Auden era mordaz respecto al conductismo: “Por supuesto, el Conductismo ‘funciona’. También la tortura. Deme un conductista con los pies en la tierra que no haga tonterías, unos cuantos medicamentos y aparatos eléctricos sencillos, y en seis meses le haré recitar el Credo de Atanasio en público” (Auden, 1971).

99. El propio Freud no ha sido superado respecto a las descripciones del psicoanálisis (véase, p. ej., Freud, 1973). Erdelyi (1985) considera la teoría freudiana desde el punto de vista de la psicología cognitiva, e informa de

evaluaciones experimentales de algunas de sus hipótesis.

99. Beck (1976, 1991) defiende que el razonamiento defectuoso conduce a neurosis, véase también Maslow (1938), Ellis (1962), Wells (2000), Cohen (2003) y Leahy (2004).

99-100. La cita es de Auden (1977, p. 395).

100. Smeets y De Jong (2005) informan de un estudio típico sobre razonamiento y enfermedad psicológica: sus resultados son equívocos.

100. La mujer preocupada por contraer el VIH es una paciente de Mancini.

100. Los test de personalidad muestran que las personas difieren en su tendencia a la neurosis, p. ej., la escala de neurosis de Eysenck (véase, p. ej., Eysenck y Eysenck, 1975), y el Inventario de Personalidad Multifásica de Minnesota [Minnesota Multiphasic Personality Inventory].

100. Los estudios sobre gemelos y su tendencia hacia el neuroticismo, o la falta de ella, aparecen en Livesley et al. (1993) y Jang et al. (1998). Es probable que estas tendencias innatas sean consecuencia de múltiples genes (cf. Ley y Ploman, 1997).

101. Oswald et al. (1970) informan de una relación entre ansiedad y dificultad para respirar en el asma y la bronquitis.

101-2. El caso de Ruth, la paciente obsesiva-compulsiva, está descrito en Leon (1990). Reed (1997) y Pélissier y O'Connor (2002), p. ej., informan sobre el razonamiento de sujetos con compulsiones obsesivas.

102. De acuerdo con la teoría del apoyo de Tversky y Koehler (1994), el describir con mayor detalle un suceso sugiere más pruebas en su favor y por tanto conduce a un aumento de la probabilidad que se le adjudica.

102. La hipótesis de fusiones en el pensamiento obsesivo-compulsivo se debe a varios autores, incluidos Wells (2000), Thodarson y Shafran (2002) y O'Connor y Aardema (2003).

103. Bruce S. Singh informa del caso de la hipocondría fugaz en su website: www.mja.com.au/public/ mentalhealth/articles/singh/singh.html#case1.

103. Salkovskis (1996) sostiene que el razonamiento en la hipocondría tiende a corroborar la perspectiva de la peor posibilidad.

103. La apuesta de Pascal aparece en sus Pensamientos (Pascal, 1966): “Vamos a sopesar la ganancia y la pérdida al apostar que Dios existe... Si ganas, ganas todo; si pierdes, no pierdes nada. Apuesta, por tanto, sin dudar que Él existe.”

104. En relación con los estudios sobre imágenes cerebrales de la ansiedad en sujetos que son obsesivos, véase Breiter et al. (1996); respecto a la ansiedad en la fobia social, véase Birbaumer et al. (1998), y en relación con la depresión, véase, p. ej., Davidson et al. (2002).

104-5. El estudio a partir de cuestionarios aparece en Johnson-Laird et al. (2006b).

105. Los experimentos sobre el transtorno obsesivo-compulsivo y la depresión se describen en Johnson- Laird et al. (2006b).

105. El test respecto a una tendencia hacia el transtorno obsesivo-compulsivo fue el Inventario de conductas obsesivo-compulsivas de Padua abreviado (Rhéaume et al., 2000), que mis colegas pasaron a 290 estudiantes de la Universidad de Palermo con el fin de seleccionar a los catorce con la puntuación más alta para la tendencia obsesivo-compulsiva y a los catorce con la puntuación más baja.

107. El test respecto a una tendencia hacia la depresión utilizó el Inventario de la Depresión de Beck (véase Beck et al., 1979), que mis colegas pasaron a una nueva muestra de 370 estudiantes en la Universidad de Palermo para seleccionar a los dieciocho participantes que tenían la puntuación más alta para la depresión y los veintidós que tenían la más baja. Un estudio no publicado de Isabelle Blanchette y Michelle Campbell mostró también una mejora inesperada en el razonamiento silogístico. Militares veteranos de la primera guerra del Golfo razonaban mejor partiendo de premisas que eran emotivas y específicas para ellos, p. ej., Algunos incidentes de fuego amigo acabaron en muerte, que partiendo de otro tipo de premisas. En un estudio similar no publicado, estas autoras evaluaron el razonamiento de sujetos de tres ciudades –Londres, Manchester y la ciudad canadiense de Londres, Ontario– respecto a los ataques terroristas con bombas de Londres. La tarea exigía que los participantes inhibieran sus creencias con el fin de alcanzar una conclusión válida. Una semana después de los ataques, los participantes de Inglaterra eran más exactos

que la muestra canadiense, pero solo con silogismos pertinentes para los ataques terroristas. No se producía diferencia con los silogismos de control o silogismos con un contenido emotivo general. Seis semanas después, la diferencia disminuía, y solo la muestra de Londres seguía teniendo ventaja sobre la muestra canadiense en los silogismos sobre terrorismo.

107. Beck (1976) discute los pensamientos de los pacientes, que él considera como “automáticos”.

107. En cuanto a descripciones de la terapia de “exposición”, véanse Andrews et al. (2003) y Tryon (2005).

Capítulo 8: Únicamente conexiones

109. El epígrafe de la Parte III pertenece al poema Carta de Año Nuevo (Auden, 1976, p. 170).

111. La famosa película Asesinato en el Oriente Express es diferente a la novela de Agatha Christie (1960).

111. El número de posibilidades con cien sospechosos es 2100, que es un número enorme: 1,267,650,600,228,229,401,496,703,205,376, es decir, más de un millón de millones de trillones.

113. George Boole (1815-1864) nació en Lincolnshire, Inglaterra, pero llegó a ser profesor de la Universidad de Galway en Irlanda. Su principal libro es Boole (1854). La introducción al álgebra booleana para niños de Mary Boole (1909) dedica un capítulo a Macbeth. El estudio psicológico de los conceptos booleanos ha dado lugar a una amplia bibliografía desde Shepard et al. (1961) hasta Feldman (2004).

115. Las tablas de verdad se describen en Jeffrey (1981). Fueron inventadas de manera independiente por Peirce y Wittgenstein.

115. En relación con las lógicas en las que no todos los teoremas válidos se pueden probar en cualquier sistema formal, véase Boolos y Jeffrey (1989). El

cálculo de predicados se puede formalizar utilizando una única regla de inferencia. Pero otra forma es utilizar reglas formales para cada conectiva. Este método se llama “deducción natural” (véanse Prawitz, 1965; Gentzen, 1969). Las teorías psicológicas sobre el razonamiento deductivo basadas en reglas formales de inferencia plasman la misma idea (véanse, p. ej., Braine y O’Brien, 1991; Rips, 1994).

115. La cita de Lord Kelvin aparece en sus clases de Baltimore (Kelvin, 1884, p. 270).

115. El programa de ordenador para el razonamiento con modelos se describe en Johnson-Laird y Savary (1999).

118. Existen varios métodos para minimizar las expresiones booleanas. El método Quine-McCluskey se describe en Quine (1955) y McCluskey (1956); otras referencias incluyen, p. ej., Gimpel (1965), Arevalo y Bredeson (1978) y Dunne (1988). La inabarcabilidad de la minimización de las expresiones booleanas se describe en Garey y Johnson (1979).

118. La idea de que nuestro sistema de razonamiento construye los modelos de uno en uno me resultó clara a partir de dos líneas de investigación independientes. Una línea son los estudios de Tom Ormerod sobre condicionales parafraseados como disyunciones y viceversa (véanse, p. ej., Richardson y Ormerod, 1997; Ormerod y Richardson, 2003). La segunda línea es la investigación de Vladimir Sloutsky que muestra que los niños tienden a ignorar la segunda cláusula de una premisa formada por una conectiva proposicional, de modo que la premisa da lugar a un modelo mental. Es más probable que la ignoren en una disyunción y menos probable en una conjunción (Morris y Sloutsky, 1998; Morris y Sloutsky, 2002). De manera similar los adultos tienden a recordar las conectivas como conjunciones (Rader y Sloutsky, 2001; Sloutsky y Goldvarg, 2004).

118. El principio de verdad se presenta en Johnson-Laird y Savary (1999). Algunos críticos han pensado que el principio significa que los modelos mentales solamente representan aquellas cláusulas que se mencionan en las premisas. Sin embargo, esta opinión implica que las oraciones tienen los mismos modelos independientemente de las conectivas que aparecen en ellas. No es así. La conectiva “y” da lugar a un modelo, mientras que la conectiva “o” da lugar a dos o tres dependiendo de si es excluyente o incluyente.

118. Johnson-Laird y Byrne (1991, 2002) sostienen que razonamos utilizando modelos mentales basados en conectivas proposicionales, pero que las notas mentales a pie de página nos permiten construir modelos completos en casos sencillos.

119. Johnson-Laird y Tagart (1969) informan sobre un estudio en el que los participantes evaluaban la verdad o falsedad de condicionales (véanse también Evans, 1972b; Oaksford y Stenning, 1992).

120. El Capítulo 2 señala que la negación no se puede representar en un icono.

121-2. Claire Walsh llevó a cabo los experimentos en los que se mostraba que los participantes sucumbían a las inferencias ilusorias basadas en disyunciones (Walsh y Johnson-Laird, 2004). Las ilusiones se producen también al pensar sobre lo que es posible (Goldvarg y Johnson-Laird, 2000). Johnson-Laird y Savary (1996, 1999) informaron de los primeros estudios sobre inferencias ilusorias. Entre los estudios que han utilizado procedimientos o instrucciones para reducir las ilusiones se incluyen Newsome y Johnson-Laird (1996), Tabossi et al. (1998), Santamaría y Johnson-Laird (2000) y Yang y Johnson-Laird (2000a,b).

123. Un artículo sin publicar de Simon Handley y Ruth Byrne de 1995 informa sobre estudios respecto a la negación de conjunciones y disyunciones (Mimeo, Universidad de Plymouth, Inglaterra). La negación de una conjunción: A y B a la vez, no, tiene tres modelos completos: no-A B, A no-B, no-A no-B. Barres y Johnson-Laird (2003) informan de que sus participantes tuvieron dificultades para imaginar qué es falso a partir de una proposición compleja.

Capítulo 9: Soy mi propio abuelo: el razonamiento sobre identidades y otras relaciones

125. Debo agradecer la historia sobre Ernest Harrison a Littlewood (1953). La teoría sobre el razonamiento con las relaciones que aparece en este capítulo está basada en Goodwin y Johnson-Laird (2005a).

125. El estudio del profesor Higgins se describe en el Capítulo 2.

126. La lógica de las relaciones llegó a dar frutos como resultado del trabajo de De Morgan, Peirce y Schröder (véase, p. ej., Kneale y Kneale, 1962). Peirce analiza la construcción de relaciones complejas, como por ejemplo, es el padre de la madre de, y su reducción a relaciones más simples (véase, p. ej., Peirce, 1931-58, Vol. 1, párrafo 66, es decir, el primer párrafo de un artículo, “The logic of relatives”, publicado originalmente en 1883). Inventa diagramas de las relaciones que anticipan las llamadas redes “semánticas” (véase Sowa, 1984) y sostiene que cualquier relación entre cualquier número de argumentos puede reducirse siempre a un conjunto de relaciones en las que ninguna relación tiene más de tres argumentos. Tarski (1965, capítulo V) da una versión moderna de las propiedades lógicas de las relaciones, tales como por ejemplo la transitividad.

126. Miller y Johnson-Laird (1976) discuten el caso de los elefantes pequeños.

126. La imposibilidad de tratar las relaciones como propiedades era conocida por los lógicos del siglo diecinueve (Winkelmann, 1980). Una relación como padre de es un tipo especial que los matemáticos conocen como una “función”. Cuando se dan valores a los argumentos de una función, esta da como resultado un solo valor único, lo mismo que la función suma da un solo valor único dados los números que hay que sumar. De igual manera, padre de George W. Bush tiene un único valor: George H. W. Bush.

127. Halford (1993) sostiene que las relaciones con más de cuatro argumentos tienden a exceder nuestra capacidad para entenderlas (véase también Halford et al., 1998). Se refiere al número de argumentos de una relación como su “complejidad relacional”. Podemos adoptar dos métodos para reducir el número de argumentos de una relación que necesitamos considerar simultáneamente. Un método reduce el número de argumentos, como sucede cuando tratamos la relación como una propiedad. El otro método divide una tarea en diferentes segmentos para evitar tener que pensar sobre relaciones complicadas (Birney y Halford, 2002).

127-8. Los marcos de referencia se discuten, por ejemplo, en Miller y Johnson-Laird (1976), O’Keefe y Nadel (1978) y Levinson (1996). El lenguaje y las partes intrínsecas de los objetos se describen en Fillmore (1971), Clark (1973) y Talmy (1983).

128. Inferencias basadas en pares de premisas como por ejemplo:

El triángulo está a la derecha del círculo.

El círculo está a la derecha del rombo.

se conocen como “silogismos lineales” o “problemas de series de tres términos”. Störring (1908) parece haber sido el primero que las estudió en el laboratorio (véase Woodworth, 1938), y Piaget (1928) también llevó a cabo algunos estudios iniciales (véase una revisión, p. ej., en Maktelow, 1999). James (1890, p. 646) propone una teoría del razonamiento sobre relaciones basada en representaciones integradas únicas. Hunter (1957) propone operaciones para construir estas representaciones. De Soto et al. (1965) sostienen que estas representaciones son imágenes visuales, y Huttenlocher (1968) muestra que encontramos más fácil añadir a una representación el sujeto que el objeto de la relación. Otras criaturas también hacen inferencias transitivas, y las palomas pueden actuar mejor que los estudiantes universitarios (véase Delius y Siemann, 1998). Clark (1969) muestra que las inferencias transitivas con un término del tipo “más ancho que” son más fáciles que las que incluyen su converso “más estrecho que”. Si la puerta es más estrecha que la mesa, la implicación es que ambas son estrechas en relación con alguna norma. Pero la proposición de que la mesa es más ancha que la puerta es neutra respecto a su tamaño en relación con la norma. Por consiguiente, términos como “más estrecho” transmiten más información que términos como “más ancho”, que a menudo corresponden al nombre de la dimensión (“anchura”). Los lingüistas se refieren a términos del tipo de “más estrecho” como marcados y a términos del tipo de “más ancho” como no marcados.

129. El uso de axiomas (“postulados del significado”) para capturar las propiedades lógicas de las relaciones se discute en Bar-Hillel (1967).

129. Las teorías de razonamiento basadas en reglas formales de inferencia se describen en el Capítulo 1.

130. Riley y Trabasso (1974), Potts y Scholz (1975) y Newstead et al. (1986) informan de los datos respecto a que los participantes construyen disposiciones espaciales. Trabasso et al. (1975) muestran que niños de seis a nueve años y

estudiantes universitarios hacen inferencias transitivas juntando elementos según una disposición ordenada. Riley (1976) replicó estos resultados y mostró que se producen con relaciones transitivas distintas a la longitud (p. ej., altura, peso, felicidad, simpatía). En Johnson-Laird y Ehrlich (1982) y Johnson-Laird y Mani (1982) se pueden encontrar descripciones de modelos de relaciones espaciales. Ruth Byrne y yo esbozamos la teoría presentada en este capítulo (Byrne y Johnson-Laird, 1989).

130. Algunos autores sugieren que utilizamos los principios lingüísticos o los modelos dependiendo del procedimiento experimental (Ormrod, 1979; Verweij et al., 1999), o como estrategias alternativas (Egan y Grimes-Farrow, 1982; Roberts, 2000). Otros autores sostienen que utilizamos ambas cosas, los principios lingüísticos y las disposiciones, en diferentes momentos del proceso (Sternberg y Weil, 1980). Hummel y Holyoak (2001, 2003) han implementado un programa "conexionista" que combina los factores lingüísticos con una disposición espacial. El que algún sistema conexionista pueda componer relaciones complejas tales como la madre del padre del hermano de sigue siendo una cuestión abierta.

131. Oberauer y Wilhelm (2000) muestran en un experimento que dada una relación espacial, como por ejemplo: la taza está a la derecha de la cuchara, es más fácil añadir el sujeto de la oración (la taza) a una representación que incluye al objeto (la cuchara) que lo contrario (véase la nota anterior respecto al trabajo de Huttenlocher).

131. Kamp y Reyle (1993) analizan las expresiones de referencia y muestran cómo se transforman en representaciones del discurso. Sus representaciones intentan ser, más que una teoría psicológica, herramientas útiles para el análisis lingüístico y por tanto no postulan representaciones icónicas, sino representaciones con una lógica transparente. Por ejemplo, en Garnham et al. (1995) y Garnham (2001) se describe una teoría psicológica respecto a cómo se interpretan las frases de referencia en términos de modelos de discurso.

132. Las estrategias inferenciales se explican en el Capítulo 19.

135. El Capítulo 2 discute los límites de la iconicidad.

136. Ruth Byrne llevó a cabo el experimento sobre el razonamiento espacial (véase Byrne y Johnson-Laird, 1989).

136. Hagert (1984) propone postulados del significado “bidimensionales”, como por ejemplo:

Para cualesquiera x, y, y z, si x está a la izquierda de y y z está delante de x entonces z está delante de x, que está a la izquierda de y.

137. Entre los estudios que corroboran que las inferencias basadas en un modelo son más fáciles que las que se basan en modelos múltiples están Schaeken et al. (1996a,b), Vandierendonck y De Vooght (1997) y Carreiras y Santamaría (1997). Podemos representar la indeterminación de la posición de un objeto dentro de un modelo o mediante una nota en un modelo (véanse Vandierendonck, Dierckx y De Vooght, 2004; Schaeken et al. , en prensa). El estudio en que todos los problemas contienen una premisa irrelevante se describe en Schaeken et al. (1998).

138. Walter Schaeken llevó a cabo el estudio sobre el razonamiento temporal basado en los tiempos verbales y los verbos auxiliares (Schaeken et al., 1996b). Los significados de las expresiones temporales son complejos; véase, p. ej., Reichenbach (1956), y Miller y Johnson-Laird (1976). Schaeken et al. (1996a) describen el programa para el razonamiento temporal.

139. Geoffrey Goodwin llevó a cabo el estudio sobre la pseudotransitividad (véase Goodwin y Johnson- Laird, 2005c).

140. La prueba de Valéry aparece en Momentos (Valéry, 1970, p. 75).

140. La respuesta al acertijo final es: ¡Yo soy mi propio abuelo! He transformado la letra (copyright Moe Jaffe y Dwight Latham, 1947) para dejar las relaciones edípicas en todo su esplendor.

Capítulo 10: Los silogismos y el razonamiento sobre propiedades

141. El cálculo lógico para cuantificadores estándar como por ejemplo las versiones idealizadas de “cualguier” y “algunos”, es decir, el cálculo cuantificativo, fue formulado por primera vez por Frege (1879/1967) y de forma independiente por Peirce (1883, véase Vol. 3, párrafo 328, en Peirce, 1931-1958). En el cálculo estándar, las variables toman a los sujetos como valores. La descripción uniforme de los cuantificadores incluidos “la mayoría” y otros que quedan fuera del cálculo cuantificativo estándar se debe a Mostowski (1957). Su defensa para el lenguaje natural aparece en Montague (1974) y Barwise y Cooper (1981). En esta lógica, el cálculo cuantificativo de “segundo orden”, las variables pueden tomar como valor tanto conjuntos de sujetos (es decir, propiedades) como sujetos.

141-2. Geurts (2003) señala la dificultad de demostrar la inferencia simple con cuantificadores estándar utilizando reglas formales de inferencia. En el Capítulo 1 se describen las teorías del razonamiento que se basan en reglas formales.

142. Bar-Hillel (1969) identifica el escándalo de que los lógicos hayan olvidado los argumentos cotidianos. Keene (1992) es una honrosa excepción.

142. En relación con estudios sobre los factores que influyen al considerar a qué cantidades se refieren las frases cuantificadas, véase Moxey y Sanford (p. ej., 1993, 2000).

144. Kant (1787/1978) argumenta en contra del uso de un único diagrama en una prueba; Wittgenstein (1953) argumenta en su favor. Hume (1739/1978) ya había mostrado cómo se podía hacer.

145-6. Hay un determinado tipo de cuantificadores que no se puede utilizar en conjunción con una premisa sobre un conjunto completo para producir una inferencia válida, concretamente, cualquier cuantificador que establezca un límite superior al número de sujetos del conjunto. Por ejemplo, la siguiente inferencia no es válida:

No más de dos de los artistas son escultores.

Todo escultor es pintor.

Por tanto, no más de dos de los artistas son pintores.

La inferencia no es válida porque los artistas que no son escultores podrían ser pintores.

146. La descripción de los silogismos de Aristóteles aparece en Primeros analíticos. La historia de la lógica de los silogismos se presenta en Kneale y Kneale (1962).

146. En relación con una descripción de la distribución en la lógica medieval, véase, p. ej., Cohen y Nagel (1934). El análisis moderno es el siguiente: Cuando los cuantificadores se tratan como relaciones entre conjuntos, en el caso de Todos los B son C, C es monotónicamente creciente, es decir, si C implica D, entonces B implica D, mientras que B es monotónicamente decreciente, es decir, si A implica B, entonces A implica C (Barwise y Cooper, 1981). El uso de modelos mentales para representar relaciones entre conjuntos fue presentado en Johnson-Laird (1983), véase también, p. ej., Neth y Johnson-Laird (1999). Geurts (2003) propone reglas formales para una teoría psicológica basada en la misma idea, debido a que son propiedades emergentes del uso de modelos y significados.

146. De manera similar a la inferencia no válida a partir de “Más de la mitad de los pintores son escultores”, Power (1984) muestra que los sujetos tienden a inferir inválidamente: “Más de una cuarta parte de los A son B; por tanto, más de una cuarta parte de los B son A”.

147. El chiste de Russell sobre el número humorístico de Mind aparece en Russell (1927). Russell (1946) señala la incapacidad de los silogismos para manejar inferencias relacionales, como por ejemplo: todos los caballos son animales; por tanto, todas las cabezas de caballos son cabezas de animales.

147. La broma de Platón sobre el perro pertenece a su diálogo sobre los sofistas, Eutidemo, 298d-298e.

147-8. Las reglas para los silogismos de la Tabla 10.2 parafrasean las de Cohen y Nagel (1934, p. 79).

149. Los experimentos sobre razonamiento silogístico se revisan en Johnson-Laird y Byrne (1991) y Evans et al. (1993). Chater y Oaksford (1999) analizan varios experimentos que examinan todas las diferentes variedades lógicas del

silogismo. Llegan a la conclusión de que los participantes tienen mayor probabilidad de hacer inferencias válidas que inferencias no inválidas.

149. La versión original de la hipótesis de la atmósfera se propone en Sells (1936) aunque se basaba en Woodworth y Sells (1935). Begg y Denny (1969) reformularon la hipótesis de la atmósfera en forma de dos principios: (1) Siempre que hay una premisa negativa, se extrae una conclusión negativa; en caso contrario, se extrae una conclusión afirmativa. (2) Siempre que hay una premisa que contiene “algunos” se extrae una conclusión que contiene “algunos”; en caso contrario, se extrae una conclusión que contiene “todos” para las premisas afirmativas y una conclusión que contiene “ninguno” para las premisas negativas. Entre los teóricos que proponen que la atmósfera forma parte del proceso de razonamiento, se incluyen García Madruga (1984), Inder (1987), Newell (1990), Polk y Newell (1995) y Rips (1994, p. 247). Revlis (1975) parece haber sido el primero en sugerir que podría ser que no razonáramos en absoluto, sino que siguiéramos los principios del efecto atmósfera. Gilhooly et al. (1993) proponen que tenemos menos probabilidad de razonar que de sacar conclusiones que encajen con la atmósfera de la más conservadora de las dos premisas (véase también Wetherick y Gilhooly, 1990). Chater y Oaksford (1999) sostienen que el razonamiento tiene sus raíces en el pensamiento probabilístico (véase también Oaksford y Chater, 1998). Dadas ciertas asunciones razonables, las conclusiones se pueden ordenar, de la que menos información contiene a la que más contiene, de la siguiente forma: Algunos A no son C, Ningún A es un C, Algunos A son C y Todos los A son C. No razonamos, sino que seleccionamos una conclusión que encaja con la forma de la premisa que menos información contiene. Su opinión es que el cálculo de probabilidades es una versión tanto normativa como descriptiva de cómo hacemos deducciones (véase el Capítulo 16).

149. Johnson-Laird y Byrne (1989) describen el estudio que tiene como cuantificador “solo”. Shaw y Johnson-Laird (1988) proporcionan otras pruebas en contra del efecto atmósfera. Entre los críticos respecto a la hipótesis de la atmósfera se incluyen Chapman y Chapman (1959), Mazzocco et al. (1974), Dickstein (1975) y Johnson-Laird y Bara (1984).

149. Los sudokus se discuten en el Capítulo 19.

149. Johnson-Laird (1983) informa de las diferencias de dificultad entre un silogismo y otro, y las diferencias en la habilidad para el razonamiento

silogístico entre una persona y otra (véase también Galotti et al., 1986). Bara et al. (1995) informan de la relación entre la memoria de trabajo y el razonamiento silogístico, y Copeland y Radvansky (2004) informan también de una correlación entre su capacidad y un razonamiento silogístico exacto. Stanovich (1999) describe los datos que relacionan inteligencia y razonamiento.

151. Bucciarelli y Johnson-Laird (1999) presentan la versión más reciente de la teoría de modelos de los silogismos y su implementación informática. Entre las descripciones previas se incluyen Johnson-Laird y Steedman (1978), Johnson-Laird y Bara (1984), que describen dos estrategias alternativas utilizando modelos, y Johnson-Laird y Byrne (1991). Los programas de ordenador posteriores eludieron las críticas técnicas (p. ej., Hardman, 1996). Una pregunta habitual sobre la teoría de modelos es qué determina qué modelo se construye en primer lugar (véase, p. ej., Chater et al., 2005, p. 304). La respuesta es el principio según el cual dada una premisa que interrelaciona, pongamos por caso, panaderos y jugadores de criquet, estando ya los autores y los panaderos representados en un modelo, los sujetos actualizan su representación de los panaderos. Cualquier relación entre autores y jugadores de cricket que no se establece de ese modo esperará a ser descubierta en la construcción de modelos posteriores.

153. Los efectos de la figura se descubrieron en un experimento que Janellen Huttenlocher y yo llevamos a cabo en 1971. Aparecen por primera vez en Wason y Johnson-Laird (1972, pp. 153-155). Dickstein (1975) informa de haber descubierto de forma independiente los efectos de la figura.

153. Oberauer y Wilhelm (2000), Oberauer et al. (2005) y Hörnig et al. (2005) informan de la explicación de Oberauer sobre los efectos de la figura y los datos en que se basa. Respecto a los efectos en el razonamiento relacional, véase Baguely y Payne (2000). También influye en los efectos un principio famoso según el cual es más fácil asignar una frase a su referente previo cuando este referente previo se menciona antes que un nuevo referente (es decir, el principio respecto a ”lo dado-lo nuevo” de Clark y Haviland, 1977).

153-4. El origen de los círculos de Euler se cuenta en Kneale y Kneale (1962). En relación con las teorías psicológicas que utilizan los círculos de Euler, véanse, p. ej., Erickson (1974), Johnson-Laird (1975), Fisher (1981) y Cardaci et al. (1996). Stenning y sus colegas tienen pruebas que confirman el uso de representaciones de posibilidades al razonar. Han mostrado una equivalencia

entre los modelos mentales y una nueva forma de razonamiento basada en los círculos de Euler. Este método elude la habitual explosión combinatoria de diagramas (véanse, p. ej., Stennint y Oberlander, 1995; Stenning y Yule, 1997). Rips (2002, p. 387) sostiene que los participantes utilizan los círculos de Euler porque los descubrieron en el colegio.

154. Edwards (2004) describe los orígenes de los diagramas de Venn y sus desarrollos modernos. Las teorías psicológicas que se relacionan con los diagramas de Venn incluyen Guyote y Sternberg (1981) y Newell (1981).

154-5. Stenning y sus colegas han formulado reglas para los círculos de Euler, reglas para los modelos mentales y reglas para las fórmulas lógicas, que dan lugar en todos los casos a las mismas inferencias válidas (p. ej., Stenning y Yule, 1997). De ahí que la forma de discriminar entre estas descripciones sea a partir de patrones de errores y otras características de la actuación.

155. En Bucciarelli y Johnson-Laird (1999) aparece un estudio en el que se dio a los participantes lápiz y papel y se les pidió que pensaran en voz alta. Ford (1995) evaluó a participantes más entendidos. Les dividió en aquellos que utilizaban círculos de Euler y aquellos que utilizaban una estrategia de sustitución verbal (descrita también en términos de modelos por Johnson-Laird y Bara, 1984). Al menos cuatro de los participantes de Ford imaginaron entidades individuales. Uno de ellos decía, por ejemplo, “... supongamos que: hay dos historiadores, lo que significa que hay dos tejedores que también son historiadores, así que podemos decir que algunos de los tejedores son historiadores...”. Desde Störring (1908) hasta Stenning (2002) los psicólogos han defendido que al razonar hay quien utiliza imágenes y quien utiliza otros procedimientos verbales.

156. El estudio de las ilusiones con disyunciones de premisas cuantificadas aparece en Yang y Johnson- Laird (2000a), y el estudio de sus antídotos en Yang y Johnson-Laird (2000b). Tanto Luca Bonatti como David O’Brien (comunicaciones personales en abril de 1997) sugirieron que la ilusiones se producen porque quienes razonan aplican incorrectamente a una disyunción de premisas, una estrategia basada en suposiciones. Esta idea es post hoc y hace predicciones erróneas sobre algunos de los problemas de control (Yang y Johnson-Laird, 2000a).

Capítulo 11: ¿No somos todos optimistas? El caso del razonamient complejo

157. En relación con una introducción al cálculo cuantificativo, véase Jeffrey (1981), quien también discute la oración, “Todo el mundo quiere a quien quiere”. Rips (1994) defiende un equivalente mental del cálculo cuantificativo.

158. Los lógicos representan en el cálculo cuantificativo las dos interpretaciones de una oración, como por ejemplo: Todas las cartas de Bury están en el mismo sitio que algunas de las cartas de Avon (véase, p. ej., Jeffrey, 1981). En relación con una comparación entre lógica y lenguaje natural, véase Johnson-Laird (1970b).

158. El orden en el que un bucle se incluye en otro dentro del programa depende del “ámbito” lógico de los cuantificadores (Jeffrey, 1981).

161. En Johnson-Laird et al. (1989) se informa sobre el programa y los experimentos que corroboran las predicciones de la teoría: las inferencias que requieren solo un modelo son más fáciles que las que requieren múltiples modelos. Greene (1992) señaló que los problemas de modelos múltiples tenían conclusiones en las que el segundo cuantificador tenía un ámbito mayor, p. ej.: ninguna de las cartas de Avon está en el mismo sitio que alguna de las cartas de Caton. Mostró que los participantes preferían extraer conclusiones de la forma: Algunas de las cartas de Caton no están en el mismo sitio que ninguna de las cartas de Avon. Victoria Shaw mostró que las inferencias de modelos múltiples eran más difíciles que las inferencias de modelo único incluso cuando todas las inferencias podían tener cualquier forma de conclusión (Shaw y Johnson-Laird, 1998).

162-3. Paolo Cherubini llevó a cabo los estudios sobre inferencias recursivas (véase Cherubini y Johnson- Laird, 2004). El estudio examinó también los problemas negativos, que eran mucho más difíciles que los problemas afirmativos. Marcin Hitczenko escribió un programa de ordenador que implementa la explicación de la teoría de modelos para estas inferencias.

164. La Tabla 11.2 se basa en un dibujo de Cherubini y Johnson-Laird (2004).

165. Mi definición de “optimismo” se hace eco de la afirmación del antropólogo Levi-Strauss según la cual cualquiera que crea que existen los bárbaros es

también un bárbaro (véase Gellner, 1970).

166. Wang (1963) discute las pruebas basadas en premisas con tres cuantificadores, como por ejemplo: “Para cada baldosa, hay al menos alguna baldosa, tal que para cualquier baldosa...” Demuestra que no existe garantía de que en este campo se pueda establecer la no validez mediante un número finito de pasos.

Capítulo 12: La modulación: un paso hacia la inducción

167. El epígrafe para la Parte IV pertenece al poema Carta de Año Nuevo (Auden, 1976, p. 162).

170. Las teorías del razonamiento basadas en la lógica formal que incluyen la regla disyuntiva son, p. ej., Johnson-Laird (1975), Rips (1983, 1994) y Braine y O’Brien (1998). La cita del difunto John Macnamara pertenece a un e-mail que me envió en 1989. Los casos que violan la regla formal: Si A entonces B; A; por tanto B; se discuten en Byrne (1989, 1991) y Byrne et al. (1999). El orden temporal puede violar la regla formal: A; B: por tanto, A y B. Por ejemplo, sería probable que fuéramos reacios a la inferencia:

Byron murió de calentura.

Cuando Byron se despertó un día se encontró con que era famoso.

Por tanto, Byron murió de calentura y cuando se despertó un día se encontró con que era famoso.

171. La inferencia sobre el ascensor se basa en un episodio de la vida del director de cine Orson Welles. Cuando se estrenaba en San Francisco su obra maestra Ciudadano Kane, entró en el ascensor de un hotel y se encontró frente a frente con el magnate de la prensa a quien él satirizaba, Willian Randolph

Hearst. Welles invitó a Hearst al estreno de la película, Hearst mantuvo un silencio gélido hasta que salió del ascensor. Welles le dijo a la imagen que se alejaba, “Charles Foster Kane hubiera aceptado”.

171. La descripción de la modulación que aparece en este capítulo se debe en gran parte a un artículo en colaboración con mi colega Ruth Byrne sobre el tema de los condicionales (Johnson-Laird y Byrne, 2002).

173. Johnson-Laird (1970b) informa de los efectos de la modulación sobre los cuantificadores. Zwaan y Radvansky (1998) ofrecen también una descripción de cómo los modelos de una situación pueden afectar a la comprensión del lenguaje.

173. Para la interpretación de los cuantificadores quasinuméricos, como por ejemplo “muchos”, véanse Moxey y Sanford (1993, 2000).

174. La dificultad para interpretar afirmaciones, como por ejemplo, “La abuela odia a cada uno de su familia”, se discute en Cherubini y Johnson-Laird (2004).

174. Tom Ormerod llevó a cabo los experimentos sobre los efectos de la modulación en las inferencias a partir de premisas disyuntivas (en colaboración conmigo). La tarea de extraer un tipo de conclusión a partir de otro tipo de premisa ha sido estudiada por varios autores (p. ej., Fillenbaum, 1976; Cheng y Holyoak, 1985; Ormerod, Manktelow y Jones, 1993; Richardson y Ormerod, 1997; Ormerod y Richardson, 2003). Bouquet y Warglien (1999) llevaron a cabo el primer estudio que mostró que las proposiciones disyuntivas que se refieren al mismo sujeto dan lugar a inferencias más fáciles que las que se refieren a dos sujetos distintos. Walsh y Johnson-Laird (2004) han corroborado este resultado.

175. Mi colega Gilbert Harman da un argumento convincente en contra de la confusión de la lógica con la psicología (véase, p. ej., Harman, 1986).

176. El Capítulo 20 describe a los teóricos que han despreciado la lógica como una herramienta analítica para el análisis de los argumentos.

176. Stenning y van Lambalgen (2005) proponen una sofisticada teoría que se sale de la lógica clásica. La interpretación del discurso es ella misma un proceso de razonamiento, dicen, y lleva, no solo a las formas lógicas de las proposiciones, que dependen del contenido y del contexto, sino también al uso de diferentes lógicas. Por ejemplo, en un tipo de interpretación, tratamos de

hacer verdaderas las premisas, mientras en otro tipo de interpretación, buscamos contraejemplos. En principio, hay infinitas lógicas posibles, y por ello Stenning y van Lambalgen se centran en cómo funciona un tipo de lógica, y cómo se puede implementar en una red de neuronas idealizadas. Opera a partir del principio, utilizado en ciertos lenguajes de programación, de que si una proposición no se puede demostrar, entonces su negación es verdadera. La lógica permite que una conclusión que se sigue de las premisas se pueda retirar a la luz de una premisa adicional. Hay similitudes entre este método y la teoría de modelos (véase el Capítulo 23). Por consiguiente, la única diferencia importante es que la modulación no necesita la selección de una lógica determinada, un asunto que sería delicado porque la modulación puede tener un número indefinido de efectos diferentes en la interpretación de las oraciones. En lugar de eso, una vez que se ha hecho una interpretación, el principio de validez semántica estándar sostiene en relación con la interpretación: una inferencia es válida si su conclusión es verdadera en todas las posibilidades compatibles con la interpretación.

176. En relación con una versión de Ramanujan y su desconocimiento de las pruebas, véase Kanigel (1991, p. 92, y véase p. 207 en relación con el citado comentario del gran matemático de Cambridge G. H. Hardy). Con un número finito de modelos finitos, siempre podemos determinar, en principio, si una inferencia es válida. Los matemáticos, sin embargo, tienen que pensar en los enteros positivos, 1, 2, 3..., y hay infinitos. Cómo razonan sobre conjuntos infinitos es un misterio que los psicólogos todavía no han aclarado.

Capítulo 13: El conocimiento y las inducciones

178. Los científicos cognitivos estudian la inducción en muy diversas áreas, a menudo independientes unas de otras, p. ej., las explicaciones sobre la conducta de otras personas (Ross, 1977), el razonamiento sobre casos legales (Ashley, 1990), el aprendizaje “basado en explicaciones” en la inteligencia artificial (Kolodner, 1993), el razonamiento causal (p. ej., Cheng, 1997), el razonamiento basado en la analogía (p. ej., Russell, 1989). La inducción está en todas partes y en ninguna.

178. Peirce (p. ej., 1955, capítulo 11) acuñó el término “abducción” o “retroducción” para referirse al paso inferencial que retorna desde una observación a una hipótesis. Sostiene que la forma de la inferencia sea:

Se observa el hecho sorprendente, C.

Pero si A fuera cierto, C sería lo lógico.

Por tanto, hay una razón para sospechar que A es verdadero.

Por tanto, pensaba en explicaciones de hechos. Mi utilización del término “abducciones” pone de relieve este punto, y permite que sean un tipo especial de inducciones. Peirce considera el razonamiento de Kepler al establecer las leyes del movimiento planetario como abductivo, porque tuvo que reconciliar las observaciones sobre las posiciones de Marte con una visión copernicana del mundo y la asunción de que el Sol “debe tener algo que ver con hacer que los planetas se muevan en sus órbitas”.

178. El Capítulo 23 describe la idea de cambios en las creencias tan pequeños como sea posible.

179. El informe oficial sobre el desastre de The Herald of Free Enterprise distribuye la responsabilidad entre todos, empezando por el director de la compañía y acabando por el ayudante del contramaestre (Sheen, 1987).

179. Logic of Scientific Discovery, de Popper (1959), descarta la inducción en ciencia.

179-80. La ley se debe a Karl Marbe, un miembro de la escuela de psicólogos de Würzburg que floreció allí antes de la Primera Guerra Mundial (Humphrey, 1951).

180. Harman (1973) y Thagard (1989) son filósofos escépticos respecto a que las explicaciones impliquen lógicamente las generalizaciones que explican.

180. Una importante idea histórica sobre la inducción de conceptos es que es una

búsqueda de lo que es común en un conjunto de entidades. Hay condiciones necesarias que consideradas en conjunto bastan para definir el concepto (véase, p. ej., Mill, 1874). Fue el primer programa para las investigaciones psicológicas en el que los participantes buscaron las condiciones necesarias y suficientes para varios conceptos, véase, p. ej., Hull (1920), que sostiene, por ejemplo, que el significado de la palabra perro es “una característica más o menos común a todos los perros y no común a los gatos, muñecas y osos de peluche”. Wittgenstein (1953) rechaza las condiciones necesarias y suficientes para los conceptos. Los juegos, dice, no tienen nada en común, solo parecidos de familia. No obstante, podría haber un juego que fuera un estereotipo. Los psicólogos ya habían tenido la misma idea, pero era un secreto bien guardado (véanse, p. ej., Fisher, 1916; Bruner et al., 1956) hasta que emergió en el trabajo de Rosch (p. ej., 1973). Un prototipo de perro, por ejemplo, tiene cuatro patas, una cola, pelo, pero estas no son condiciones necesarias –un perro podría tener tres patas, ser calvo y mudo–. No todos los ejemplos de un concepto son igualmente representativos –un terrier es un prototipo de perro, pero un chihuahua no lo es–. De manera similar, las opiniones sobre pertenencia a un concepto dependen de la distancia entre el ejemplo y el prototipo (véanse, p. ej., Rips et al., 1973; Hampton, 1979). Minsky en su noción de valor por defecto aclaró la lógica de los prototipos. Un perro tiene cuatro patas, por defecto, pero si un perro determinado tiene solo tres patas, entonces esta información anula el valor por defecto sin contradicción. La lógica por defecto subyace en los llamados lenguajes de programación “orientada a objetos”, como C++, Java y Lisp. Los “objetos”, como las ventanas en las pantallas de nuestros ordenadores, tienen ciertas propiedades por defecto.

180. Barsalou (1983) fue un precursor en el estudio de los conceptos ad hoc, como por ejemplo: cosas que llevar a un picnic.

181. McFarlane (1984) describe el descubrimiento de la penicilina por Fleming.

181. Nisbett et al. (1983) muestran que los participantes son sensibles a las probabilidades de las propiedades cuando hacen inducciones. Osherson et al. (1986) discuten otras descripciones de la inducción. Describen algunos principios generales que parece ser que seguimos cuando hacemos inducciones sobre las propiedades de las entidades.

181. Mill (1874, p. 226) remite a la inducción por enumeración, una noción que tiene su origen en Novum Organon de Francis Bacon (véase Bacon, 2003).

181. Los experimentos sin publicar en los que los participantes extraen inducciones específicas fueron llevados a cabo en colaboración con Tony Anderson.

182. Quienes razonan de forma prudente evalúan sus generalizaciones inductivas, y una amplia literatura psicológica ha mostrado que tienden a centrarse en ejemplos positivos de sus hipótesis (véanse, p. ej., Wason, 1960; Klayman y Ha, 1987).

182. Un guion representa una secuencia de acciones estereotipadas y por defecto, como por ejemplo pedir un café en un bar (p. ej., Bower et al., 1979; Schank y Abelson, 1977).

183. Los teóricos han diseñado muchos programas de ordenador que simulan la inducción de conceptos (p. ej., Hunt et al., 1966; Michalski, 1983; Langley et al., 1987). Otros han implementado programas que establecen representaciones "conexionistas" de conceptos (p. ej., Paccanaro y Hinton, 2000). Polya (1973) propone que las reglas formales son las bases heurísticas de las inducciones y muchos teóricos posteriores se basan en ellas, p. ej., Collins y Michalski (1989). Las reglas formales para la inducción de la Tabla 13.1 pertenecen a Winston (1975), Hayes-Roth y McDermott (1978), Michalski (1983) y Thagard y Holyoak (1985).

184. Los postulados del significado se discuten en el Capítulo 9.

184. Las dificultades para utilizar la visión en el reconocimiento de objetos se discuten, p. ej., en Varr (1982) y Ullman (1996).

184. Hume (1739/1978) sostiene que no existe justificación racional para la inducción. Goodman (1983) plantea la pregunta escéptica sobre cómo distinguir entre aquellas propiedades que apoyan las generalizaciones inductivas y aquellas que no lo hacen.

184. Yo sostengo que dada una hipótesis que descarta una parte de las posibilidades, su número de generalizaciones inductivas aumenta exponencialmente con el número de conceptos o proposiciones que podrían ser importantes (véase Johnson-Laird, 1993).

184. Kahneman y Tversky (1973) descubrieron que la disponibilidad de conocimiento es uno de los principales heurísticos que subyacen en las opiniones

sobre las probabilidades de los sucesos.

184. Berwick (1986) propone la restricción de que las inducciones deberían ser tan “informativas” como fuera posible, y la obtiene a partir de un teorema de la teoría del aprendizaje formal que se debe a Angluin (1978).

185. Winston (1975) implementó la idea de un profesor servicial, que establece una secuencia adecuada de ejemplos que presentar a los alumnos, de forma que ellos puedan inducir conceptos, como por ejemplo el de “arco”.

185. Goodman (1972) identifica algunos de los problemas filosóficos de la similitud. Tversky (1977) contiene una descripción del concepto que ha tenido gran influencia. La similitud depende a menudo de una teoría del dominio (véanse, p. ej., Murphy y Medin, 1985; Rips y Collins, 1993). Depende también del contexto y de otra serie de factores (Medin et al., 1993). Los niños sintonizan enseguida con la idea de que los murciélagos tienen características esenciales que, a pesar de las apariencias, les distinguen de los pájaros (p. ej., Gelman y Markman, 1986; Keil, 1989; Markman, 1989).

185. Cualquier buen libro de texto de estadística discute la regresión a la media. Kahneman y Tversky (1973) informan de que tenemos dificultades con el concepto, y a menudo no lo tenemos en cuenta como explicación sencilla de los hechos –un fallo que se encuentra en la bibliografía psicológica.

186. En el Capítulo 5 se discute el pensamiento “mágico”.

Capítulo 14: El método de Sherlock Holmes: la abducción

187. La descripción que hace Conan Doyle del Dr. Joseph Bell como inspiración para Sherlock Holmes aparece en Sebeok y Umiker-Sebeok (1983).

187. Las deducciones de Holmes sobre el francmasón aparecen en La aventura de la liga de los pelirrojos (Conan Doyle, 1930).

188. La afirmación de Holmes de que sus conclusiones eran tan infalibles como Euclides aparece en Un estudio en escarlata, capítulo 2 (Conan Doyle, 1930, p.

23).

188. Respecto a la cita de La aventura de la caja de cartón, véase Conan Doyle (1930, p. 895).

188. Peirce (1955) describe la abducción.

188. El capítulo 22 analiza las relaciones causales.

188-9. La inferencia para la mejor explicación se discute, p. ej., en Peirce (1931-1958, 5.180), Harman (1965) y Lipton (1991).

189. Para tener una descripción de la navaja de Ockham, véase Kneale y Kneale (1962, p. 243). Kolmogorov defiende una medida de la complejidad mínima de un objeto en función de la menor longitud de las instrucciones para que una herramienta computacional sencilla describa el objeto (véase, p. ej., Li y Vitanyi, 1993). Como muestra Chaitin (1998), una vez que los programas exceden una longitud bastante pequeña, no hay ningún programa del que podamos garantizar que prueba que un programa es tan corto como es posible. La parsimonia es un objetivo excelente, pero imposible de probar.

189. La idea de que una explicación debería tener un máximo alcance, incluida la predicción de nuevos sucesos, se discute, p. ej., en Lipton (1991) y Howson y Urbach (1993).

189. Miyake (1986) pidió a parejas de participantes que explicaran cómo funciona una máquina de coser, y observó que tendían a explicar el mecanismo a un nivel, y luego a turnarse explicándolo en un nivel inferior, etc.

189. Rozenblit y Keil (2002) describen la “ilusión” de la profundidad explicativa.

190. El hecho de que los occidentales atribuyan las acciones al estado mental del que actúa se conoce como “error fundamental de atribución” (véase, p. ej., Ross, 1977). Morris et al. (1955) dan argumentos a favor de diferencias en el razonamiento de los occidentales y los orientales.

190-1. Johnson-Laird et al. (2004a) informan del experimento en el que los participantes preferían abducir una explicación causal a deducirla. En un estudio sin publicar, Louis Lee ha mostrado que tanto los participantes occidentales en

Princeton como los participantes chinos en Hong Kong prefieren construir explicaciones abductivas en lugar de deducciones, y que es más probable en ambos casos que ofrezcan una explicación deductiva si han llevado a cabo una tarea deductiva en la misma sesión experimental.

192. El programa para construir cadenas causales se describe en Johnson-Laird et al. (2004). Fenker et al. (2005) presentan pruebas de que almacenamos relaciones causales en forma de parejas causa-efecto.

193. La explicación respecto al anuncio dirigido a los hombres pelirrojos se debe encontrar, por supuesto, en La aventura de la liga de los pelirrojos (Conan Doyle, 1930). Si no quiere usted estropear su disfrute de la historia, por favor, no lea la nota siguiente.

193. El anuncio es parte de un plan para sacar de su casa únicamente a un determinado hombre pelirrojo, de forma que desde el sótano de su casa un ladrón pueda excavar un túnel hasta un banco contiguo.

193. Tony Anderson llevó a cabo el experimento sin publicar en el que los participantes explicaban pares de oraciones aleatorias.

193. Oaksford y Chater (1991, 1998) sostienen que el razonamiento cotidiano depende de un procedimiento abarcable. La idea central de mi argumento en el texto es que hasta que no tengamos una versión clara de qué es lo que se supone que se produce como resultado del razonamiento cotidiano, no podemos decidir este asunto. Ragni (2003) probó que el razonamiento espacial bidimensional es inabarcable. Si el razonamiento espacial forma parte del razonamiento cotidiano entonces el razonamiento cotidiano también es inabarcable.

194. Tony Anderson llevó a cabo el experimento relacionado con el asesinato en el cine.

195. Un factor crítico en la construcción de un modelo es, como establecieron Tversky y Kahneman (1973, 1974), la disponibilidad de conocimiento relevante.

195-6. Las tres principales teorías psicológicas sobre la analogía se deben a Gentner (1983), Keane (1988) y Holyoak y Thagard (1989). Falkenhainer et al. (1989), Keane (1988) y Holyoak y Thagar (1989) describen programas de ordenador que implementan sus respectivas teorías. Gick y Holyoak (1983) informan de un estudio experimental que muestra que los participantes no

siempre captan el significado de una analogía. Chalmers et al. (1992) sostienen que estas descripciones eran insatisfactorias porque estaban demasiado alejadas de la percepción. Esta observación es correcta, pero la percepción da lugar a modelos mentales, que considero la base de la analogía (cf. Sowa y Majumdar, 2003, que describen un programa para extraer analogías utilizando una representación gráfica).

196-7. La analogía de Wright entre la caja y el biplano está descrita en Jakab (1990).

197. Feynman (1988, p. 54) informa de su incapacidad para poner en palabras la forma de un cigüeñal.

197. La cita es de Feynman (1985).

198. Si nuestras razones para las acciones son también causas es un enigma que discuten los filósofos (véase, p. ej., Davidson, 1980).

198. El aforismo es de Lichtenberg (1958).

198. Los comentarios de James Watson sobre los hechos y la teoría se citan en Crick (1988, p. 60).

198. Karmiloff-Smith e Inhelder (1974/75) señalan que los niños inicialmente ignoran los contraejemplos a sus hipótesis.

Capítulo 15: El balance de probabilidades

199. Franklin (2001) sostiene que el cálculo de probabilidades podría no haberse inventado hasta el siglo diecisiete porque las herramientas matemáticas esenciales eran desconocidas hasta esa fecha.

199. Las diversas interpretaciones de las probabilidades incluyen frecuencias (p. ej., von Mises, 1957), implicaciones lógicas parciales (p. ej., Keynes, 1943) y grados de creencia (p. ej., Savage, 1954). Cada una tiene sus defensores, y no hay consenso.

200. Los estudios de Kahneman y Tversky sobre el uso de heurísticos en las opiniones intuitivas sobre probabilidad están resumidos en un conjunto de artículos en Kahneman et al. (1982). Esta colección incluye el artículo de Tversky y Kahneman (1973) sobre el uso del conocimiento accesible, y el artículo de Tversky y Kahneman (1982a) sobre el heurístico de "representatividad", que incluye el estudio con el ejemplo de Linda (véase también Kahneman y Tversky, 1972). La opinión de que una conjunción, como por ejemplo feminista y cajera de banco, es más probable que uno de sus constituyentes, como por ejemplo cajera de banco, se conoce como la "falacia de la conjunción" (Tversky y Kahneman, 1983). De acuerdo con la "teoría de apoyo" de Tversky y Koehler (1994), la descripción de un suceso con mayor detalle da lugar a más pruebas en su favor y por ello conduce a atribuirle una mayor probabilidad (véase también Miyamoto et al., 1995). Kahneman y Frederick (p. ej., 2005) contrastan nuestras intuiciones inductivas con los juicios más medidos de nuestro pensamiento deliberado. Gilbert (1989) hace algo similar.

201. El problema de las tres puertas procede de un programa de televisión americano cuyo presentador era Monty Hall (cf. Falk, 1992; Nickerson, 1996). La respuesta correcta se describe al final del capítulo. El autor de la columna "Pregunta a Marilyn" en la revista Parade publicó un análisis correcto del problema. Gracias al poder de equiprobabilidad, como señala Falk (1992, p. 203), Marilyn recibió miles de cartas de los lectores, muchos de ellos de universidades e institutos de investigación, y aproximadamente el noventa por ciento insistían en que estaba equivocada.

201. La Corte Británica de Apelaciones expresó su escepticismo sobre el uso del teorema de Bayes en estos términos:

Se permitió a la defensa deducir pruebas del Teorema de Bayes en conexión con la evaluación estadística del perfil de ADN. Aunque sus Señorías expresaron opiniones no concluyentes sobre el asunto, tenían muy graves dudas sobre si esa prueba era adecuadamente admisible... sus Señorías nunca habían oído sugerir que un jurado debiera considerar la relación entre ese tipo de prueba científica y otra prueba en relación con la fórmula de probabilidad...

Citado de The Times (de Londres), 9 de mayo, 1966, informando sobre la

opinión de la Corte de Apelación en Regina v. Adams.

201. La opinión de que los seres humanos poseen un módulo inferencial innato que plasma el cálculo de probabilidades se debe a Cosmides y Tooby (1966; véase también Gigerenzer y Hoffrage, 1995). En los experimentos de Cosmides y Tooby, los participantes fueron más exactos en los problemas definidos con frecuencias que en los definidos con porcentajes de probabilidades. Cosmides y Tooby escriben (p. 62): “Los mecanismos basados en frecuencias no podrían hacer surgir razonamiento bayesiano a menos que nuestras mentes contuvieran mecanismos que plasman al menos algunos aspectos del cálculo de probabilidades.” Howson y Urbach (1993, p. 422) expresaban escepticismo respecto a esta conclusión. De la misma manera, los sujetos novatos pueden razonar sobre problemas en relación con las posibilidades de sucesos únicos, pero el uso de frecuencias no garantiza el razonamiento satisfactorio (véase Girotto y Gonzalez, 2001, 2002, 2006; Evans et al., 2000; y Sloman et al., 2003). Los psicólogos (y los probabilistas) discuten sobre si los asertos sobre probabilidades de sucesos únicos son sensatos. Con un aserto, como por ejemplo: “La probabilidad de que la Tercera Guerra Mundial empiece mañana es ,01”, el problema es especificar las condiciones en las que sería cierto. No obstante, somos felices asignando probabilidades a sucesos únicos, así que no parece que los encontremos sin sentido (para encontrar argumentos opuestos, véanse Kahneman y Tversky, 1996, quienes aceptan esta opinión, y Gigerenzer, 1996, que no lo hace).

201. Gordon (2004) describe experimentos con una pequeña tribu amazónica, los Pirahä, que muestran que su sistema de contar es, “aproximadamente uno”, “aproximadamente dos” y “muchos”. Es difícil imaginar que tengan un módulo innato para razonar sobre frecuencias (véase también Pica et al., 2004, en relación con resultados similares para otro grupo amazónico, los Mundurukú). Este segundo grupo tiene una comprensión intuitiva de conceptos geométricos, p. ej., pueden escoger el único triángulo que no contiene un ángulo recto en un conjunto de seis (véase Dehaene et al., 2006).

201-2. En relación con la teoría del razonamiento sobre probabilidades utilizando modelos mentales, véase Johnson-Laird et al. (1999).

202. El principio de “indiferencia” o “razón insuficiente” tiene una larga historia

en la teoría de probabilidades (véase Hacking, 1975). Puede llevar a dificultades (Howson y Urbach, 1993, p. 52), aunque si la indiferencia se aplica a la representación de problemas y a los modelos mentales en particular, las dificultades se pueden resolver (véase Johnson-Laird et al., 1999). El principio está en el centro de la conocida computación de Laplace (1951) según la cual las apuestas respecto a que mañana saldrá el Sol serían de 1.826.214 contra 1.

202. El experimento sobre las canicas dentro de la caja se describe en Johnson-Laird et al. (1999). Estos autores informan también de estudios sobre inferencias ilusorias respecto a probabilidades relativas. Johnson-Laird y Savary (1996) muestran que, a partir de ciertas premisas, los participantes infieren como el más probable de dos sucesos uno que, en realidad, es imposible, y a partir de otras premisas, infieren como el menos probable de dos sucesos uno que, en realidad, es cierto.

202. Bar-Hillel y Falk (1982) estudiaron el problema sobre los dos hijos. Contiene una serie de sutilezas que ha aclarado Nickerson (1996).

204. Stevenson y Over (1995) fueron los primeros que postularon el etiquetado de modelos mentales con probabilidades numéricas (véase también Stevenson y Over, 2001).

204. El experimento en el que los participantes estimaban probabilidades condicionales conversas pertenece a un estudio sin publicar que Jennifer L. Adams llevó a cabo como estudiante sin graduar en la Universidad de Princeton. Hammerton (1997) muestra que los novatos que razonan pueden no darse cuenta de que un problema no establece un índice base esencial.

205. El principio de subconjuntos se describe en Johnson-Laird et al. (1999).

208. En relación con las pruebas en contra de que utilicemos el teorema de Bayes, véase, p. ej., Falk (1992). Shimojo y Ichikawa (1989) informan también de que incluso cuando a la gente se le enseña el teorema de Bayes siguen considerándolo contrario a la intuición.

210. Kahneman y Tversky (1973) informan sobre el estudio del índice base y la estimación de si los sujetos eran abogados o ingenieros.

210. Hacking (1975) es el filósofo que escribe sobre los dados en tiempos de los romanos.

Capítulo 16: Los contraejemplos

213. El epígrafe para la Parte V pertenece al poema The History of Truth (Auden, 1976, p. 463).

215. La idea de los juegos en la comprensión de los cuantificadores debo agradecerla, en parte, a la idea de juegos lingüísticos de Wittggenstein (1953). Hintikka (1973) es su principal exponente.

216. Holyoak y Glass (1975) informan de un estudio sobre contraejemplos para generalizaciones.

216. Las referencias a la tarea de selección aparecen en las notas al Capítulo 1. Respecto a la opinión de que no es una valoración del razonamiento deductivo, véase Oaksford y Chater (1998). Sostienen que en vez de eso estamos tratando de optimizar la cantidad de información que esperamos ganar al volver una carta. Esta opinión combina la teoría estadística de la información de Shannon (Shannon y Weaver, 1949) con el teorema de Bayes para el cálculo de probabilidades (véase el Capítulo 15). El primero postula que cuánto más improbable es una señal, más información comunica - con una concomitante reducción de la incertidumbre por parte del receptor de la señal. El teorema de Bayes nos permite computar la probabilidad de una hipótesis a la luz de algunos datos. Oaksford y Chater sostienen que al evaluar un condicional de la forma si A entonces B, en muchas de las versiones de la tarea, es racional no elegir la carta no-B. Sin embargo, los participantes dejaban de hacerlo incluso cuando podían llegar a ganar dinero a partir de una evaluación correcta de la verdad o falsedad del condicional, como Carlos Santamaría, Orlando Espino y yo mostramos en un experimento sin publicar. Asimismo, Stanovich (1999) informa de que los participantes más inteligentes tenían más probabilidad de seleccionar contraejemplos en la tarea de selección (véase también Stanovich y West, 1998).

216. En relación con las pruebas que apoyan la hipótesis de que la disponibilidad de contraejemplos mejora el comportamiento en la tarea de selección, véanse Love y Kessler (1995), y Liberman y Klar (1996). Las instrucciones para comprobar violaciones también mejoran el comportamiento en la tarea abstracta (Platt y Griggs, 1993; Green, 1995, 1997; Green y Larking, 1995; Griggs, 1995;

Dominowski, 1995). Sperber et al. (1995) utilizan un procedimiento más indirecto para hacer que los contraejemplos sean más relevantes - en el sentido de Sperber y Wilson (1986), y mejorar de ese modo la actuación.

216. La revista que rechaza publicar ningún artículo más sobre la tarea de selección es Cognition.

216. Wason (1960) describe su tarea 2-4-6. Klayman y Ha (1987) sostienen que los participantes no están tratando de confirmar sus hipótesis, sino que, en vez de eso, están comprobando ejemplos positivos de dichas hipótesis. Existe una amplia bibliografía sobre la tarea, véase, p. ej., Manktelow (1999) para encontrar una revisión.

216. El principio de verdad se presenta en el Capítulo 6.

217-9. Patricia Barres llevó a cabo los experimentos en los que los participantes tratan de imaginar contraejemplos para varias proposiciones (véase Barres y Johnson-Laird, 2003).

219-20. La teorías del razonamiento basadas en reglas formales se describen en el Capítulo 1. Quine (1974) y Barwise (1993) señalan que si no conseguimos encontrar una prueba formal para una inferencia, no podemos saber con certeza que no es válida: podemos haber pasado por alto alguna prueba. Por el contrario, un contraejemplo es una demostración manifiesta de no validez. Chater et al. (2005, p. 302) relacionan la aceptación de una conclusión para la que no se ha encontrado un contraejemplo con un procedimiento de un programa de ordenador conocido como “negación como fallo”: una proposición se acepta si no puede probarse su negación.

220. La descripción de las respuestas a los silogismos de las personas que razonan pobremente es de Johnson-Laird (1983, p. 120).

220. Polk y Newell (1995) proponen una teoría de los silogismos, basada en modelos, que menosprecia el uso de contraejemplos. Escriben (p. 553) : “La cuestión es que los datos de los silogismos pueden explicarse sin postular una estrategia de falsificación, en conformidad con la hipótesis de que tal estrategia de razonamiento específico es menos importante que los procesos verbales para explicar la deducción”.

220-1. Monica Bucciarelli llevó a cabo los experimentos en los que los

participantes utilizaban modelos externos de las premisas (Bucciarelli y Johnson-Laird, 1999). En contraste con los resultados de Bucciarelli, Newstead et al. (1999) informan de un estudio en el que el número de diagramas alternativos que dibujaron los participantes no predijo su exactitud en el razonamiento silogístico. Sin embargo, Newstead et al. (2002) encontraron que una medida de la diferencia entre los participantes, que reflejaba su habilidad para generar posibilidades alternativas, correlacionaba con su capacidad para refutar conclusiones silogísticas no válidas. De hecho, Newstead et al. (1999) informan de que menos de una quinta parte de las respuestas en su Experimento 3 fueron correctas. Roberts (2006) señala que los estudios sobre silogismos infravaloran la utilización de los contraejemplos. El descubrimiento de que construimos modelos externos para extraer conclusiones silogísticas muestra que estamos intentando hacer deducciones (con todo respeto a Chater y Oaksford, 1999).

221. Martín-Cordero y González-Labra (1994) son los críticos que sostienen que la teoría de modelos no consigue especificar cómo buscamos los contraejemplos.

222. Uri Hasson llevó a cabo los experimentos sobre estrategias en el rechazo de inferencias no válidas basadas en conectivas proposicionales (véase Johnson-Laird y Hasson, 2003).

223. Hansjoerg Neth llevó a cabo el experimento sobre el razonamiento con el cuantificador “más de la mitad” (véase Neth y Johnson-Laird, 1999). Este cuantificador no se puede capturar en el cálculo cuantificativo estándar, pero se puede manejar como una relación entre conjuntos (Mostowski, 1957).

227. James Kroger llevó a cabo el experimento sin publicar sobre las imágenes cerebrales y los contraejemplos en colaboración con mi colega John Cohen y yo mismo. Debido a Goel (p. ej., Goel y Dolan, 2004) y mi colega Dan Osherson (p. ej., Osherson et al., 1998; Parsons y Osherson, 2001) existe una creciente bibliografía sobre imágenes cerebrales y razonamiento. No es fácil dar sentido a los diversos resultados, pero muestran que la deducción implica regiones del cerebro que no forman parte de los procesos lingüísticos normales, y que activa regiones en el hemisferio derecho (con todo respeto a Wharton y Grafman, 1998).

229. En estudios sin publicar, Michelle Cowey y Ruth Byrne han mostrado que los participantes actúan mejor en la tarea 2-4-6 al refutar las hipótesis de los

otros que las suyas propias y que los maestros de ajedrez son mejores que los novatos al refutar los movimientos planeados por ellos.

229-30. Oaksford y Chater (1998) siguen a Anderson (1990) al suponer que el conocimiento humano es una adaptación racional al mundo (véase también Chater et al. , 2005), y sus opiniones se presagian en alguna medida en la propuesta de Evans, Over y sus colegas según la cual las probabilidades desempeñan un papel en el razonamiento (p. ej., Stevenson y Over, 1995, 2001; Evans y Over, 1996, 2004). Ellos aplican su método a tres tipos principales de razonamiento, la tarea de selección, el razonamiento simple a partir de premisas condicionales y el razonamiento silogístico (Oaksford y Chater, 2001). Sostienen: “las personas razonan mucho mejor sobre probabilidades de lo que Tversky y Kahneman suponían” (Oaksford y Chater, 1998, p. 262). En realidad, cometemos errores sistemáticos en el razonamiento probabilístico (véase el Capítulo 15). Sostienen que su explicación probabilística considera qué se está computado, mientras que los procesos que lo computan deberían ser abarcables y podrían utilizar “heurísticos rápidos y frugales” (Gigerenzer y Goldstein, 1996) o redes bayesianas (Pearl, 1988, 2000). Schroyens y Schaeken (2003) informan de que sus metaanálisis de los estudios que aparecen en la bibliografía (Schroyens et al., 2001a,b) muestran que la teoría de modelos proporciona una descripción más exacta del razonamiento condicional que la teoría probabilística. Espino et al. informan de datos sobre movimiento ocular que apoyan la teoría de modelos en lugar de la teoría probabilística para el razonamiento silogístico (véase también Espino et al., 2000b).

230. En el Capítulo 3 se informa de los datos sobre la diferencia entre conclusiones posibles y necesarias (véase Bell y Johnson-Laird, 1998; Evans et al., 1999). Markovits y Handley (2005) informan del estudio sobre razonamiento que partiendo de las mismas premisas comparaba las instrucciones para evaluar si las conclusiones eran necesariamente ciertas con las instrucciones para evaluar las probabilidades de que fueran ciertas. Los estudios muestran que los sujetos con una amplia memoria de trabajo tienen más tendencia a hacer deducciones silogísticas correctas que aquellos con una memoria de trabajo menor (p. ej., Gilhooly et al., 1999; Copeland y Radvansky, 2004). Verschueren et al. muestran que en el razonamiento a partir de premisas condicionales, los sujetos con una amplia memoria de trabajo tienen más probabilidad de utilizar contraejemplos, mientras que los que tienen una baja capacidad de memoria de trabajo tienden a basarse en consideraciones probabilísticas (véase Markovits et al., 2002, para encontrar resultados compatibles). Como sugieren Verschueren et al. , la

distinción sugiere que las consideraciones probabilísticas dependen del razonamiento heurístico intuitivo, mientras que la utilización de contraejemplos depende de procesos analíticos deliberados. Un factor pertinente es el papel del conocimiento –la modulación, dicho con mis palabras–, y Politzer y Bourmaud (2002) revisan sus efectos. Para encontrar una evaluación general del método de Oaksford y Chater, véase Rips (2002), y para encontrar una crítica a su descripción del razonamiento silogístico, que señala que no ofrece una explicación para la utilización espontánea de diagramas, véase Roberts y Sykes (2005).

230. Beth (1971) sostiene que el principio fundamental de la validez semántica es que un argumento es válido si su conclusión debe ser verdadera en el caso de que sus premisas sean verdaderas. En lógica existen métodos tanto formales como semánticos para buscar contraejemplos (véase Jeffrey, 1981). El método formal se basa en el llamado método del "árbol". El procedimiento estándar empieza con una lista de las premisas y la negación de la conclusión atribuida. Estas oraciones solo pueden ser todas verdaderas si hay un contraejemplo. Y por tanto las reglas de inferencia se formulan de forma que permiten una búsqueda sistemática de las posibilidades. El método del árbol, sin embargo, también se puede formular como un método que proporciona conclusiones de una forma directa –un método que sería la versión de la lógica más cercana a la teoría de modelos mentales–. Sin embargo, ninguno de los que proponen reglas formales en psicología ha propuesto una teoría de este tipo.

230. La correlación entre habilidad para el razonamiento e inteligencia se debe a Stanovich (p. ej., 1999). Varios estudios informan de la correlación entre habilidad para el razonamiento y capacidad de la memoria de trabajo (véase, p. ej., Bara et al., 1995; Barrouillet y Leças, 1999; véase también Kyllonen y Christal, 1990).

230. Evans y Over (1996) distinguen entre racionalidad1 para el razonamiento cotidiano y racionalidad2 para el razonamiento deductivo deliberado. Shafir (1994) ha demostrado violaciones de la transitividad en las preferencias de las abejas.

Capítulo 17: Verdades, mentiras y razonamiento superior

233. Al principio de su ensayo sobre la verdad, Bacon escribe: “¿Qué es la verdad?, dijo el burlón Pilatos y no esperaba una respuesta” (véase Bacon, 1597/1985).

233. Tarski (1944) distingue metalenguaje y lenguaje objeto, y describe la paradoja que resulta de que el lenguaje natural sea su propio metalenguaje. La oración sobre “la nieve es blanca” pertenece también al artículo de Tarski. La teoría de la “correspondencia” de la verdad es defendida también por Wittgenstein (1922) y muchos otros. En relación con una descripción reciente de los análisis filosóficos de la verdad, véase Blackburn (2005).

233. El Capítulo1 distingue entre oraciones que tienen significados y proposiciones que tienen significación.

234. Keats escribe sobre el “razonamiento consecutivo” en una carta a su amigo Benjamin Bailey (véase Gittings, 1970, p. 36), una carta que también muestra que Keats creía que lo que la imaginación valora como belleza debe ser verdadero (cf. su oda a una urna griega).

234. Rascar (1973) discute la teoría de la coherencia de la verdad. La teoría ha interesado a los científicos cognitivos que estudian el mantenimiento de las creencias consistentes (véase el Capítulo 23).

234. Grife (1975, 1989) analiza las convenciones que gobiernan el discurso cotidiano, entre las que incluye el principio de decir la verdad.

234. Medvedev (1990) describe el trágico episodio de los ingenieros de Chernobyl.

234. En relación con una descripción de la verdad en lógica, véase, p. ej., Boolos y Jeffrey (1989).

235. La definición de negación de Aristóteles aparece en De Interpretatione (véase Aristóteles, 1984, Vol. 1). En relación con una descripción de la negación en el lenguaje, véase Horn (1989).

235. La idea de que las oraciones pueden no conseguir expresar proposiciones a menos que sus presuposiciones sean ciertas se remonta a Austin (1970) y a la

crítica de Strawson (1950) sobre la teoría de las descripciones definidas de Russell, tales como por ejemplo “el actual rey de Francia” (Russell, 1905).

235-6. La idea de encajar un modelo en otro se debe a Kamp y Reyle (1993).

236. La paradoja del “mentiroso” se analiza, p. ej., en Barwise y Etchemendy (1987).

237-9. El lógico Raymon Smullyan tiene casi el monopolio de los problemas de caballeros y villanos (p. ej., Smullyan, 1978). Como él muestra, están relacionados con el teorema de Gödel según el cual existen verdades en aritmética que no pueden ser demostradas en ningún sistema formal consistente. El problema de la Tabla 17.1 se basa en uno de Smullyan (1978, p. 22).

237-9. Rips (1989) informa del estudio sobre problemas de caballeros y villanos. Su reto a los teóricos de los modelos era: “Produzcan una versión explícita del razonamiento sobre problemas de caballeros y villanos, es decir, (a) explícita desde el punto de vista teórico, (b) adecuada desde el punto de vista empírico y (c) que no sea una simple variante en cuanto a la notación de la teoría de deducción natural” (p. 113). En relación con una respuesta a este reto, véase Johnson-Laird y Byrne (1990), que describe una implementación para ordenador de cinco estrategias basadas en modelos para los problemas de caballeros y villanos.

240. Byrne et al. (1995) informan del estudio sobre los problemas de caballeros y villanos que muestra que los participantes desarrollan diferentes estrategias. Byrne y Handley (1997) corroboran estos resultados. Schroyens y sus colegas extienden la teoría de modelos hasta explicar la naturaleza de los errores en estos problemas (véanse Schroyens, 1997; Schroyens et al., 1996, 1999).

241. Johnson-Laird y Byrne (1991) describen problemas metalógicos.

242. Goffman (1959) discute las inferencias en contextos sociales.

242. En relación con las descripciones del problema de los tres sombreros, véanse, p. ej., Fujimara (1978), Anno y Nozaki (1984) y Dewdney (1989). George Erdos, de la Universidad de Newcastle en el Reino Unido, ha mostrado en un experimento sin publicar que el problema de dos sombreros es más fácil que el de tres sombreros. Erdos informa también de que los sujetos deducen con frecuencia que Anne no puede ver dos sombreros negros, pero cuando llegan a

considerar la situación de Beth olvidan que Beth también puede hacer esa deducción.

243. En relación con las transcripciones de las cintas sobre las discusiones de la Casa Blanca durante la crisis de los misiles con Cuba que se han editado y la cita sobre Mr. McCone, véase May y Zelikow (2001, p. 61).

244. Dixit y Nalebuff (1991) ofrecen una divertida introducción a la teoría matemática de juegos, incluido el teorema de Nash. Reinhard Selten hizo su escéptico comentario sobre la teoría en el Taller Internacional sobre Juegos y Modelos Mentales (Universidad de Venecia, agosto 1998). Eugenia Goldvarg-Steingold llevó a cabo el estudio sin publicar en el que sujetos novatos fueron preparados para participar en juegos experimentales sin conocer los beneficios del otro jugador.

245. Mazzocco et al. (1974) examinan la capacidad de sus sujetos para proporcionar premisas perdidas en silogismos. Sus resultados están en contra de la hipótesis de la "atmósfera" (véase el Capítulo 10).

246. Victoria Shaw llevó a cabo un estudio sin publicar en el que los participantes iniciaron el desarrollo de principios formales que gobiernan las inferencias silogísticas. Vadeboncoeur y Markovits (1999) también informan de que sus sujetos aludían a conceptos formales.

Capítulo 18: Sobre el desarrollo

247. El epígrafe de la Parte VI pertenece a un ensayo sobre Walter de la Mare (Auden, 1973, p. 389).

249. Fodor (1980) sostiene que todos los conceptos, las conectivas lógicas y los cuantificadores son innatos, y que no pueden descomponerse en primitivos. El problema con su argumento es que también parece descartar la posibilidad de que los conceptos podrían haber evolucionado dado que los procesos evolucionistas se pueden imitar en programas de aprendizaje (véase Johnson-Laird, 1983, pp. 142- 144). La evolución según la síntesis neo-Darwinista se imita en procesos "genéticos" debidos a Holland (véase, p. ej., Holland et al.,

1986; Mitchell, 1998). Dado que los humanos tienen solo 30.000 genes, todos ellos tendrían que codificar conceptos en vista de que son innatos y no se pueden descomponer en conceptos más primitivos.

249. El concepto de la mente como una pizarra en blanco se remonta a los filósofos empíricos británicos, especialmente Locke (1979), aunque en realidad él hablaba de una hoja de papel en blanco. Pinker (2002) es uno de los críticos principales respecto a este punto de vista.

249. En el libro se describen los programas conexionistas (véase el Glosario). Shastri y Ajjanagadde (1993) presentan un sistema conexionista para el razonamiento sencillo. El problema es explicar la composición de los significados (Fodor y Pylyshyn, 1988) y el razonamiento con oraciones más complejas. Los conexionistas han respondido a esta crítica (p. ej., Smolensky, 1990; pero cf. Fodor y McLaughlin, 1990). Un reto sencillo para los conexionistas es diseñar un sistema que aprendiera a hacer inferencias utilizando tablas de verdad en el cálculo proposicional. En mi opinión, solo se puede superar el reto dividiendo la tarea en módulos independientes. Falmagne (1980) defiende el aprendizaje convencional como una forma de mejorar la capacidad de razonamiento.

250. Cook y Newson (1996) es una introducción a la descripción de la gramática universal de Chomsky (véase también Chomsky, 1995).

250. Pinker (1997) defiende el valor adaptativo del lenguaje. En relación con las opiniones de Chomsky y sus colegas sobre la evolución del lenguaje, véanse, p. ej., Hauser et al. (2002) y Fitch et al. (2005).

250-1. Piaget (1896-1980) fue la encarnación de los procesos formales en el razonamiento (véase, p. ej., Inhelder y Piaget, 1958). En relación con las descripciones de la “conservación” piagetiana de las cantidades, véase, p. ej., Piaget e Inhelder (1969). Respecto a la descripción del desarrollo que aparece en la Tabla 18.1, véase Inhelder y Piaget (1964, p. ej., pp. 292-293); en relación con la equivalencia entre operaciones formales y cálculo proposicional, véase Inhelder y Piaget (1958, p. 305), y respecto a las relaciones entre las estructuras de los niños y las estructuras matemáticas, véase Beth y Piaget (1966). El colectivo matemático Bourbaki, se describe en Halmos (1957). Respecto a las exégesis de la teoría de Piaget, véanse, p. ej., Flavell (1963) y Boden (1980); y respecto a una visión neopiagetiana, véase Karmiloff-Smith (1992). Una

profética observación piagetiana: los niños tienden a no representar posibilidades múltiples (Piaget, 1987).

251-2. Braine y Rumain (1983), que defienden las reglas formales de inferencia, escriben: “La lógica de Piaget no se puede desarrollar en la adolescencia ni en ningún momento: es demasiado problemática para mantenerse como un modelo psicológico de nada.”

252. El concepto de potencia computacional se describe en el Capítulo 3.

253. Baillargeon et al. (1985), Baillargeon (1987), Baillargeon y Hanko-Summes (1990) y Spelke (1990, 1991), por ejemplo, informan sobre los experimentos que muestran que los bebés tienen el concepto de que los objetos siguen existiendo cuando están fuera de la vista y que desarrollan pronto una física “ingenua”.

254. El estudio sobre el razonamiento disyuntivo de los niños aparece en un artículo sin publicar de Justin Halberda: “Is this a dax which I see before me? Use of the logical argument disjunctive syllogism supports word-learning in children and adults.”

254. El perro de Crisipo se describe, p. ej., en Annas y Barnes (1985, p. 36). Lamentablemente, Watson et al. (2001) no consiguieron corroborar que los perros puedan hacer la inferencia. Mandler (2004) informa sobre el estudio respecto a las inducciones de los bebés.

255. En relación con las versiones sobre la “teoría” de la mente de los niños y los estudios sobre falsas creencias, véanse, p. ej., Leslie (1987, 1990) y Premack (1990).

255. Los fenómenos de la dificultad de los niños con las falsas creencias de segundo orden son sólidos (p. ej., Wimmer y Perner, 1983; Perner et al., 1987; Perner, 1991; y para una revisión, véase Wellman et al., 2001). Las potenciales explicaciones para los fenómenos son alucinantes. Respecto a la opinión de que la carga de memoria es el factor decisivo, véase Bloom y German (2000), y en relación con la opinión de que nuestro conocimiento es una maldición y no podemos recapturar nuestra antigua ignorancia, véase Birch (2005), quien informa de efectos similares sobre el razonamiento de los adultos (p. ej., Keysar et al., 2003). Respecto a la opinión de que la capacidad para enfrentarse a las falsas creencias depende de una capacidad para representar posibilidades “contrafactuales”, véanse, p. ej., Riggs et al. (1998) y Peterson y Riggs (1999).

Harris (2000) en su extraordinario libro sobre la imaginación de los niños (y los adultos) defiende que nuestra capacidad para imaginar posibilidades alternativas y calcular sus implicaciones surge de forma temprana en nuestra infancia y dura toda la vida. Pero al enfrentarse con falsas creencias puede haber mucho más en juego que el hecho de prever posibilidades contrafactuales. Perner et al. (2004) informan de que las dos cosas no van necesariamente de la mano: los niños pueden responder a preguntas sobre hechos contrafactuales antes de poder responder preguntas sobre falsas creencias incluso cuando ambas están basadas en el mismo guion. Otra posibilidad es que en la raíz del problema esté la dificultad para cambiar la atención de una alternativa a otra (véanse los resultados sin publicar de Aisling Murray y Ruth Byrne).

256. Respecto a la opinión de que los niños autistas carecen de una "teoría" de la mente, véase, p. ej., Baron-Cohen et al. (1985). Leslie y Ghaiss (1992) muestran que los niños autistas pueden enfrentarse con fotografías que no son verídicas.

256. Case (1985a,b) propone que el aumento en la capacidad de la memoria de trabajo capacita a los niños para realizar tareas más complejas, p. ej., poner las varillas en el orden correcto según la serie.

256. Chi (1976) muestra que la capacidad de la memoria de trabajo de los niños aumenta a medida que se hacen mayores. Muestra también que este aumento es en parte un resultado del conocimiento. Por ejemplo, los adultos pueden tener presentes el doble de formas sin sentido que los niños, pero con los nombres que son muy familiares para los niños, como por ejemplo "perro", la ventaja de los adultos es mucho menor (véase también Chi, 1978).

257. Fodor (1983) establece de manera radical la hipótesis de que el cerebro tiene módulos independientes.

257. Cosmides (1989) es el psicólogo evolucionista que propone un módulo innato para el razonamiento sobre el engaño. Fiddick (2004) defiende además un módulo para tomar precauciones contra el riesgo. Pinker (1997) presenta una defensa de la psicología evolucionista y discute el papel pintado floreado y las disputas entre hermanos. Entre los críticos que señalan explicaciones alternativas para los resultados con la tarea de selección se incluyen Cheng y Holyoak (1989) y Buller (2005).

257-8. Entre los críticos con la psicología evolucionista se incluyen también

Simon (1991b), que establece la analogía con las historias “simples”, y Lewontin (1990), que señala la dificultad para evaluar el método. Chomsky en una conversación prestó atención al problema de cómo el módulo adecuado desencadena una situación.

258. Una respuesta aritmética al problema del lazo de cuerda: un cuadrado de 5 cm de lado tiene un perímetro de 20 cm y un área de 25 cm2. Un rectángulo que tenga 8x2 cm sigue teniendo el mismo perímetro pero un área de solo 16 cm2. Debo agradecer este ejemplo al difunto A. R. Jonckeere.

258. Dasen (1994) informa sobre el estudio respecto al razonamiento espacial de los aborígenes australianos.

258. Keil (1989) describe el desarrollo de la capacidad de los niños para establecer categorías de objetos. Vosniadou y Brewer (1992), Samarapungavan et al. (1996) y Dicheva y Close (1996) informan respectivamente de estudios sobre los modelos mentales que tienen los niños respecto a la Tierra, el sistema solar y la recursión en computación.

258-9. Carey (1985a,b) sostiene que los niños están equipados con una o dos teorías intuitivas, que incluyen una física ingenua. Asemeja los cambios en los conceptos de los niños con cambios de una teoría científica a otra (véase Kuhn, 1970), e informa del diálogo con su hija en Carey (1988). El desarrollo, como subraya Carey (1991), implica que los conceptos de los niños pueden parecer no comparables con los de los adultos, y que “... pueden surgir nuevos conceptos que no se pueden definir en función de conceptos sostenidos anteriormente”.

259. La teoría del color de Newton aparece en su óptica (Newton, 1952). Goethe (1970) presenta su teoría. Algunos pintores criticaron la teoría de Newton porque confundían la mezcla de la luz en una rueda de colores con la mezcla de pigmentos en una paleta. Ribe y Steinle (2005) defienden la utilización de experimentos exploratorios por parte de Goethe.

259-60. El difunto Paolo Bozzi durante los años cincuenta del siglo veinte en Italia (véase la descripción en su autobiografía, Bozzi, 1991) y McCloskey (1983) en Estados Unidos fueron pioneros en el estudio de la física ingenua. Kaiser et al. informan sobre el estudio de las bolas o el agua que salen de tubos curvados.

260. Aristóteles analiza la rapidez en Física (232a, líneas 25-29).

260. El diálogo sobre la rapidez se basa en Feynman et al. (1963).

261. Barrouillet y sus colegas han llevado a cabo muchos estudios en los que niños y adultos enumeran las posibilidades para los condicionales (véanse, p. ej., Barrouillet y Leças, 1998, 1999; Barrouillet et al., 2000). La interpretación "lógica" de los condicionales se describe en el Capítulo 8.

262. Johnson-Laird et al. (1986) y Bara et al. (1995) informan de los estudios sobre el razonamiento de los niños con silogismos.

262. En Mehler et al. (1988) se informa de un estudio que muestra que los bebés de solo unos días atienden a la lengua de su madre.

262. Curtiss (1977) informa de un trágico caso moderno de "niño salvaje" –un niño que creció sin toparse con ningún lenguaje.

Capítulo 19: Estrategias y culturas

263. Las inferencias ilusorias se describen en los Capítulos 8 y 15, y el metarrazonamiento se discute en el Capítulo 17.

264. Wood (1969) y Wood et al. (1974) informan de los estudios pioneros sobre el razonamiento relacional. Estos estudios demuestran el uso de atajos. Ormrod (1979) muestra que la pregunta de un problema limita las estrategias. Quinton y Fellows (1975) describen cinco estrategias, incluyendo aquella en la que los participantes intentan resolver la pregunta únicamente a partir de la información contenida en la primera premisa. Los participantes desarrollan también estrategias para razonar sobre parentesco, p. ej., hermano del padre ser refiere al mismo familiar que tío (p. ej., Wood y Shotter, 1973). Algunas equivalencias, incluido el ejemplo anterior, es probable que estén almacenadas en la memoria a largo plazo, pero otras tenemos que deducirlas.

264. Rips (1989, 1990, 1994) es el más destacado entre los que proponen la opinión de que el razonamiento depende de una única estrategia determinista. Por el contrario, la teoría de modelos postula una diversidad de estrategias de razonamiento (véase, p. ej., Johnson-Laird, 1983; Johnson-Laird y Bara, 1984;

Byrne y Handley, 1992).

264. Girotto et al. (1977) muestran que el orden de las premisas afecta a la dificultad de las inferencias con la forma: Si A entonces B, no-B, ¿qué se sigue? Cuando aparece primero el condicional, la memoria de trabajo ya está preocupada durante la interpretación de la premisa categórica, y por tanto es improbable que quien razona construya un modelo adicional completo. Pero, cuando la premisa categórica aparece primero, capacita a quien razona para rechazar inmediatamente el modelo mental explícito.

265. Mis colegas y yo utilizamos el término “estrategia” en el mismo sentido de Bruner et al. (1956). Es también similar a la idea de “plan” de Miller et al. (1960, p. 16). Sternberg (p. ej., 1977, 1983, 1984) analiza los componentes de las analogías, los problemas de series numéricas y los silogismos. Hunt (1999, 2002) trata la mente como un procesador de información. Mis colegas y yo seguimos esta línea.

265. Una buena fuente de información respecto a la historia de los Sudoku es la enciclopedia on-line Wikipedia: en.wikipedia.org/wiki/Sudoku.

265. En el cuento de hadas de Hans Andersen (1996) La reina de las nieves, la tarea de Kay es resolver un rompecabezas que forma la palabra “Eternidad”.

265. La solución al Sudoku de la Figura 19.1 es:

7 9 1 8 4 3 2 6 5

8 5 4 6 7 2 1 3 9

3 6 2 5 1 9 7 4 8

5 4 3 9 8 1 6 2 7

6 7 8 2 5 4 3 9 1

2 1 9 3 6 7 8 5 4

4 2 7 1 9 6 5 8 3

9 3 5 7 2 8 4 1 6

1 8 6 4 3 5 9 7 2

266. Nuestra competencia al razonar con cuantificadores en los Sudoku es contraria a las opiniones según las cuales los participantes novatos desde el punto de vista de la lógica no hacen deducciones ni siquiera con premisas con un único cuantificador (con todo respeto a, por ej., Revlis, 1975; Gilhooly et al., 1993; Chater y Oaksford, 1999). También es contraria a la afirmación: “las personas solo hacen un esfuerzo en el razonamiento deductivo cuando se les instruye explícitamente para hacerlo” (Evans y Over, 1996, p. 158). Oaksford y Chater (p. ej., 1998) sostienen que la validez deductiva es un criterio inadecuado para nuestro razonamiento (véase mi discusión sobre sus opiniones en el Capítulo 16).

266. En relación con la afirmación sobre los orientales, véase la nota posterior.

266. Louis Lee llevó a cabo los experimentos con los Sudoku, y la investigación la realizó en colaboración con Geoff Goodwin (véase Lee et al., 2006).

267. La complejidad relacional es el número de argumentos, o limitaciones, en una relación (véase Halford, 1993; Halford et al., 1998).

268. El problema de la Tabla 19.1 pertenece a un experimento de Van der Henst et al. (2002), quienes describe las cinco estrategias de la Tabla 19.2.

271. El modelo nulo se describe en el Capítulo 8. Evans (1972a) informa del primer estudio psicológico sobre reductio ad absurdum, utilizando inferencias similares a estas.

271-2. La utilización de suposiciones en las teorías de reglas formales es más limitada que en la teoría de modelos. Rips (1994) permite que se hagan suposiciones solo trabajando marcha atrás a partir de una conclusión putativa; Braine y O’Brien (1998) permiten que se hagan suposiciones solo después de que las reglas de inferencia directas se hayan agotado. Nuestros resultados son contrarios a ambas versiones.

272. La descripción del desarrollo de estrategias aparece en Van der Henst et al.

(2002).

273. Van der Henst et al. (2002) informan de los experimentos sobre los sesgos de las estrategias en el razonamiento proposicional.

274. El informe sobre quienes apoyan respectivamente a Kerry y a Bush procede del Programa sobre Actitudes Políticas Internacionales [Program on International Policy Attitudes].

274. Geoffrey Goodwin llevó a cabo los experimentos respecto a cómo desarrollan los participantes estrategias para razonar sobre relaciones entre relaciones (véase Goodwin y Johnson-Laird, 2005b).

274. Los informes introspectivos pueden ser erróneos respecto a por qué los participantes actúan como lo hacen (Nisbett y Wilson, 1977). Schooler et al. (1993) interrumpieron a las personas que estaban tratando de resolver problemas que dependían de un insight. Después de la interrupción, aquellos que tenían que hacer un informe retrospectivo sobre su pensamiento resolvían menos problemas que los que llevaban a cabo una tarea que no tenía relación (un crucigrama). Otros estudios, sin embargo, informan de que el pensar en voz alta mejora la actuación (p. ej., Berardi-Colecta et al., 1995). Puede que retarde a la persona, pero con frecuencia no parece tener más efectos importantes (cf. Russo et al., 1989). En general, puede ser una guía fiable de la secuencia de pensamientos de un sujeto (Newell y Simon, 1972; Ericsson y Simon, 1980). Bell (1999) comparó el razonamiento cuando los participantes piensan en voz alta y cuando no lo hacen. Obtuvo los mismos patrones medios de resultados para ambas situaciones.

275. Los relativistas sostienen que ningún conjunto de convenciones lógicas es universal en todas las culturas (p. ej., Barnes y Bloor, 1982) o en todas las épocas (Burke, 1986). A Geertz (1984) no le gusta el relativismo pero lo prefiere a su alternativa, el racionalismo.

275. Evans-Pritchard (1937) describe a los Azande.

276. El psicólogo ruso A. R. Luria (1934, 1976) informa de los primeros estudios sobre diferencias interculturales en el razonamiento de participantes sin escolarizar en Uzbekistán. Cole y sus colegas plantearon problemas de razonamiento a los Kpelle (Cole et al., 1971). Scribner (1977) sugiere que los participantes sin escolarizar adoptan una aptitud “empírica” frente a los

problemas de razonamiento, es decir, utilizan su conocimiento y experiencia.

276. Harris y sus colegas informan sobre el estudio en Recife (Dias et al., 2005; véase también Harris, 2000).

276. Nisbett (2003) resume su descripción de las diferencias en el razonamiento entre los occidentales y los orientales (véase también, p. ej., Peng y Nisbett, 1999). Norenzayam et al. (2002) describen sus experimentos sobre las diferencias culturales en el razonamiento. Informan de un experimento que mostró que los coreanos cometen más errores que los estudiantes norteamericanos al evaluar la validez deductiva de los argumentos. Un análisis diferente de sus datos hecho por Unsworth y Medin (2005) mostró, por el contrario, que los occidentales no eran mejores que los coreanos en esta tarea.

276-7. Louis Lee llevó a cabo el estudio sin publicar en el que se presentaron problemas inferenciales ambiguos tanto a estudiantes chinos de Hong Kong como a estudiantes de Princeton. También llevó a cabo otros varios estudios sin publicar en los que se comparan muestras de estas dos poblaciones. Hasta ahora no ha encontrado ninguna diferencia fiable en su capacidad para hacer deducciones.

277. Nisbett (2003, p. 170) escribe: “La diferencias entre los dos estarían en que los americanos simplemente están más acostumbrados que los coreanos a aplicar reglas lógicas (sic) a los sucesos ordinarios...” Más adelante, comenta, “... realmente no encontramos que los orientales tengan problemas con la lógica formal...” (p. 188).

278. En relación con una cultura que no tiene números, véase la nota a Gordon (2004) en el Capítulo 15.

278. En el Capítulo 2 describo la navegación de los habitantes de Micronesia (véase, p. ej., Gladwin, 1970). Tulkin y Konner (1973) describen la caza en el Kalahari y muchos otros antecedentes de pensamiento científico. Kroeber (1961) describe cómo un indígena americano corta una punta de flecha a partir de una botella de coca-cola.

Capítulo 20: Cómo podemos mejorar nuestro razonamiento

279. La rapidez de procesamiento correlaciona con la capacidad de razonamiento y la actuación en test de inteligencia. A medida que uno se hace mayor, se vuelve más lento, pero, como Patrick Rabbitt y sus colegas han mostrado, lo que afecta a la rapidez es la atrofia del cerebro en la vejez. Los análisis muestran que la edad, que representa a otros cambios neurofisiológicos, también tiene un efecto sobre la inteligencia. Por tanto, la rapidez de procesamiento –o la falta de ella– da cuenta de gran parte de la lentitud intelectual en la tercera edad, pero no de todas las diferencias de inteligencia entre una persona anciana y otra (para encontrar una valiosa revisión, véase Rabbitt y Anderson, 2006).

279. Nuestra inteligencia a los 11 años predice en parte cuánto vamos a vivir. Las medidas de inteligencia también predicen el comportamiento en la vida académica y en el trabajo. Algunos especialistas dicen que el "CI", que se basa en test de inteligencia, es la mejor forma de predecir actualmente muchos aspectos de nuestras vidas (en relación con una revisión de estos resultados, véase Deary, 2005).

279. Northcutt (2000) es el primer volumen de Los premios Darwin.

279. En relación con el informe sobre el desastre del barco The Herald of Free Enterprise, véase Sheen (1987).

280. El proyecto Venezuela fue iniciado en 1979 por el entonces presidente. El Dr. Machado, ministro para el Desarrollo de la Inteligencia Humana, tenía el objetivo de elevar el nivel de inteligencia de toda la población. Entre los expertos a los que consultó sobre el proyecto estaban Jaacov Agam, Berry Brazelton, Edward de Bono, Reuven Feuerstein, David Perkins y Robert Sternberg. El proyecto finalizó en 1984 junto con el gobierno que lo instituyó. ¿Funcionó? Tuvo un efecto sobre el rendimiento en los colegios. Pero una respuesta adecuada a la pregunta exigiría el mismo esfuerzo que se puso inicialmente en el proyecto. Para encontrar una descripción del proyecto, véase, p. ej., www.newhorizons.org/trans/international/dickinson_venezuela.htm.

280. Los lógicos tienden a asumir que la lógica es buena para nosotros. Por ejemplo, Hintikka en el resumen de una charla pronunciada en la Conferencia Internacional sobre Razonamiento Práctico Formal y Aplicado en Bonn en 1996

escribe: “... la lógica (deductiva) es estratégicamente hablando una llave para todo buen razonamiento”.

280. Lowenstein (2000) es una divertida versión de cómo los economistas ganadores de un premio Nobel perdieron una cantidad prodigiosa de dinero dirigiendo un fondo de reserva.

280. Debo agradecer el argumento de Paine (1989, pp. 94-95) sobre la religión a Keene (1992), que es el mejor libro que conozco sobre el análisis lógico de argumentos reales. Mi análisis se diferencia un poco del de Keene y es mucho menos formal.

282. Barwise y Etchemendy (1992, 1994) describen Hyperproof. Stenning (2002) informa de sus estudios empíricos investigando su eficacia como herramienta para enseñar lógica. Y Sternberg (1997) describe once estilos diferentes de pensamiento, incluido uno gráfico.

282. Cheng et al. (198) informan sobre el estudio en el que los participantes recibían un curso de lógica de un semestre antes de tratar de resolver la tarea de selección de Wason.

282. Los sucesos producidos en el reactor nuclear de Three Mile Island se describen en el Capítulo 3.

284. Lin Yutang escribe, “Me parece que modernos economistas y psicólogos tienen una sobredosis de escrupulosidad y no el suficiente insight. Quizá este es un punto sobre el que no se puede hacer suficiente hincapié, el peligro de aplicar la lógica a los asuntos humanos” (véase Lin, 1937, pp. 424-431, citado por Keene, 1992, p. 5).

284. La bibliografía sobre la lógica informal ha aumentado a pasos acelerados a partir de las críticas de Toulmin (1958) y Scriven (1976) respecto a la lógica. Entre los teóricos que defienden la identificación de falacias se incluyen Kahane (1984), Govier (1988) y Johnson y Blair (1994). Algunos de los que abogan por la utilización de diagramas especiales para analizar los argumentos son Beardsley (1959), Freeman (1988) y Fisher (1988). Walton (1989) defiende un análisis pragmático del discurso. Tanto Toulmin (1958) como McPeck (1981) afirman que los principios del razonamiento difieren de un campo de experiencia a otro. Los sistemas de ordenador para analizar argumentos se describen en Verheij (2003). Entre los estudios psicológicos sobre la argumentación informal

se incluyen Perkins et al. (1990), Kuhn (1991), Voss et al. (1993), Rips (1998), Rips et al. (1999) y Green y McCloy (2003). El patrón de resultados de estos estudios es complejo, pero han mostrado que el rendimiento es pobre y que a menudo quienes razonan pasan por alto líneas alternativas de discusión. Asimismo, muchas de las estrategias que surgen en los estudios de laboratorio sobre inferencias deductivas reaparecen en el razonamiento informal (Kuhn, 1991, p. 274).

284. Las opiniones filosóficas sobre la utilización de diagramas en pruebas geométricas se discuten en Beth (1971).

284. Lewis Carroll (1958) describe su juego de lógica. Peirce ideó dos sistemas independientes de diagramas para el razonamiento lógico (véanse las notas al Capítulo 2).

284. Simon (1991a, p. 96) informa de que los ingenieros encontraban más fáciles de entender los casos de la Corte Suprema cuando utilizaba diagramas de circuitos en los que las posiciones de encendido se correspondían con las decisiones de la Corte. Los experimentos de Simon con diagramas se publicaron en Larkin y Simon (1987) y Tabachneck y Simon (1992). Larkin y Simon (1987, p. 71) escriben: “En vista de los efectos dramáticos que pueden producir las representaciones alternativas en la búsqueda y reconocimiento de procesos, puede parecer sorprendente que los efectos diferenciales sobre la inferencia parezcan menos sólidos. La inferencia es en gran medida independiente de la representación, si el contenido de la información de los dos conjuntos de reglas de inferencia [una que trabaja con diagramas y otra que trabaja con definiciones verbales] es equivalente...”

284. Barwise y Etchemendy (1992, p. 80) escriben: “Es mucho más difícil utilizarlos [los diagramas] para presentar información imprecisa, información negativa o información disyuntiva.”

285-7. Las inferencias basadas en dos premisas disyuntivas se describen en el Capítulo 3. Malcolm Bauer llevó a cabo los experimentos sobre la utilización de diagramas en el razonamiento a partir de premisas disyuntivas (véase Bauer y Johnson-Laird, 1993). Un estudio de Boudreau y Pigeau (2001) muestra que el razonamiento espacial también es más fácil cuando los participantes razonan a partir de diagramas que cuando lo hacen a partir de descripciones.

287. El capítulo anterior describe la estrategia de razonamiento en la que los participantes dibujaban diagramas para seguir el rastro de posibilidades independientes compatibles con las premisas.

288-91. Victoria Bell (1999) ideó el método de modelos y llevó a cabo los experimentos para investigar su eficacia.

Capítulo 21: Los misterios del si

293. El epígrafe de la Parte VII pertenece a Auden (1976, p. 231).

295. El físico y filósofo escéptico del siglo segundo antes de Cristo Sextus Empíricus en Adversus mathematicos cita el epigrama de Calímaco (véase Kneale y Kneale, 1962, p. 128).

295. Medvedev (1990) describe lo que fue mal en Chernobyl. En relación con una revisión de la gran dificultad de la inferencia a la que se enfrentaron allí los ingenieros, véase, p. ej., Evans et al. (1993).

296. Entre los libros recientes sobre condicionales están Eells y Skyrms (1994), Levi (1996), Woods (1997), Lycan (2001), Bennet (2003), y Gauker (2005). Evans y Over (2004) revisan tres importantes tipos de teorías psicológicas sobre los condicionales, que discutiré posteriormente en este capítulo.

296. Alma, la hija de Vittorio Girotto, es la niña de tres años que dio una hábil respuesta a su promesa sobre las palomitas.

297. La teoría de los condicionales presentada en este capítulo es resultado de la colaboración con Ruth Byrne (véase, p. ej., Johnson-Laird y Byrne, 2002).

297. En relación con un análisis de la gramática de los condicionales, véase Dudman (1988).

297-8. Una excepción menor a la afirmación de que las cláusulas sí describen posibilidades es la aparición de preguntas en ellas. Pueden plantear preguntas del tipo: quién, qué, dónde, siempre que la cláusula repita la forma de un comentario

anterior, p. ej:

Si llegara Vivien entonces se iría Evelyn.

Evelyn se iría si llegara ¿quién?

298. Grife (1975) defiende la interpretación lógica de los condicionales, reforzada con varias herramientas pragmáticas.

298-9. Barrouillet y su colega han llevado a cabo los principales estudios en los que los participantes enumeran las posibilidades compatibles con los condicionales (véase, p. ej., Barrouillet, 1997; Barrouillet y Leças, 1998, 1999). Sus resultados apoyan la tendencia relacionada con el desarrollo, en cuanto a la enumeración de posibilidades, y una medida de la capacidad de la memoria de trabajo correlacionaba con el número de posibilidades que enumeraba un niño, y seguía haciéndolo incluso cuando en el análisis no se consideraba el nivel escolar. Otros resultados son compatibles con esta tendencia relacionada con el desarrollo (véase, p. ej., Taplin, Staudenmayer y Taddonio, 1974; Russell, 1987; Markovits, 1993). Con frecuencia los niños pequeños interpretan los condicionales como similares a conjunciones, es decir, como compatibles con una única posibilidad (véase, p. ej., Paris, 1973; Delval y Riviere, 1975; Kuhn, 1977; Politzer, 1986).

298-9. Johnson-Laird y Tagart (1969) y Tagart y Evans (1972b) informan sobre los estudios en los que los participantes evaluaban la verdad o falsedad de condicionales. Kneale y Kneale (1962, p. 549) discuten la historia de las “paradojas”.

300. Evans y Over (2004) sostienen que el razonamiento condicional se basa en la suposición de que la cláusula si es cierta. Escriben: “Los variados y numerosos usos de ‘si’ en el lenguaje cotidiano, que hemos mostrado a lo largo de este libro, tienen todos algo en común: se invita a quien escucha a suponer o imaginar una situación hipotética” (pp. 171-172).

300. Evans y Over (2004, pp. 19-20) sostienen que las paradojas son inferencias absurdas. Rechazan las teorías que las consideran válidas.

300-1. En Byrne y Tasso (1999), Johnson-Laird y Byrne (2002) y en una serie de estudios llevados a cabo por Cristina Quelhas (véase, p. ej., Quelhas y Byrne, 2003) se informa de experimentos en los que los participantes enumeran posibilidades para diferentes tipos de condicionales.

301-2. Los experimentos sobre inclusiones y exclusiones espaciales pertenecen a Johnson-Laird y Byrne (2002). Otros conocidos patrones de inferencia se pueden explicar mediante la modulación, p. ej., la tendencia de los sujetos a hacer inferencias de la forma: Si A entonces B, no A, por tanto , no B. Si podemos pensar en otras formas de provocar B con independencia de A, entonces es probable que no hagamos la inferencia (como han mostrado muchos estudios, véase, p. ej., Markovits y Vachon, 1990; Cummins et al., 1991; Thompson, 1994; Cummins, 1995; Janveau-Brennan y Markovits, 1999). Markovits y sus colegas, incluido Barrouillet, han desarrollado una explicación para este fenómeno basada en modelos (Markovits et al., 1998; Markovits y Barrouillet, 2002; Markovits y Quinn, 2002; Quinn y Markovits, 2002).

302-3. La interpretación natural de las oraciones del “burrito” [“donkey” sentences] según la lógica no funciona (véase, p. ej., G. Evans, 1980). Por tanto, esta interpretación es errónea: Si existen un x tal que x es un burro y Joe es propietario de x, entonces Joe pega a x. El ámbito del cuantificador existe un x está restringido a la cláusula si del condicional, porque no existe garantía de que tal x exista. Por tanto, el x en la cláusula entonces no está “ligado” [“bound”] al cuantificador. Una solución es representar la oración con un cuantificador universal (véase, p. ej., Reinhart, 1986): Para cualquier x, si x es un burro y Joe es dueño de x, entonces Joe pega a x.

303. En relación con otros análisis del misterio de la cerilla mojada, véase, p. ej., Stalnaker (1968), Lewis (1973), Lycan (1991) y Over y Evans (1997).

303-4. Minsky (1975) introduce las asunciones por defecto. Investigadores en inteligencia artificial (IA) posteriores han diseñado muchos tipos de sistemas formales que utilizan reglas o axiomas para acomodar el razonamiento por defecto, y otros sistemas que permiten extraer conclusiones a la luz de premisas posteriores. Esta rama de la IA se conoce como razonamiento “no monotónico” (en relación con una revisión, véase Brewska et al., 1997).

305. Barres y Johnson-Laird (2003) informan sobre el estudio de las conjunciones y disyunciones de condicionales.

306. La nota a pie de página de Ramsey aparece en la p. 155 de Ramsey (1990), un ensayo sobre la causalidad publicado por primera vez en 1929.

306. En el capítulo siguiente esbozo un mecanismo para la construcción de cadenas causales.

306. Las teorías filosóficas sobre el significado de los condicionales que son generalizaciones del test de Ramsey se formulan dentro del marco de "mundos posibles" (véase, p. ej., Kripke, 1963). Un "mundo posible" es una situación posible, si bien una que es una especificación completa, y el mundo real se considera un miembro del conjunto de mundos posibles. Según el análisis de Stalnaker (1968), la verdad de un condicional contrafactual en un mundo, el mundo 1, depende de otro mundo, el mundo 2, que es accesible al mundo 1 y que es el que más se parece al mundo 1 excepto en que en él se cumple la cláusula si del condicional. Si la cláusula entonces es verdadera en el mundo 2, entonces el condicional como un todo es verdadero; en otro caso, es falso. No es obvio que exista siempre un único mundo totalmente similar al mundo real excepto en que una cláusula si contrafactual sea cierta. Si existe siempre un único mundo de este tipo entonces una de estas dos proposiciones es cierta y una de ellas es falsa:

Si Verdi y Bizet hubieran sido compatriotas entonces los dos habrían sido italianos.

Si Verdi y Bizet hubieran sido compatriotas entonces los dos habrían sido franceses.

Aun así, es imposible decidir qué proposición es la verdadera. A la luz de este tipo de ejemplo, Lewis (1973) sostiene que un condicional es verdadero solo en caso de que la cláusula si y la cláusula entonces sean verdaderas en un mundo que está más cerca del mundo real que cualquier mundo en el que la cláusula si es verdadera y la cláusula entonces falsa.

307. La cita de Max Hastings pertenece a Armageddon, su historia sobre los últimos ocho meses de la Segunda Guerra Mundial (Hastings, 2004, p. 43).

307-8. Byrne (2005) reúne su trabajo y otras investigaciones sobre los condicionales contrafactuales y sobre cómo pensamos respecto a las alternativas a la realidad. Byrne y Tasso (1999) y Quelhas y Byrne (2003) informan de los experimentos sobre la inferencia de Chernobyl.

308. Quine (1952) propone la teoría según la cual un condicional no tiene valor de verdad cuando su cláusula si es falsa, Wason (1966) y Johnson-Laird y Tagart (1969) la aceptan, y Evans y Over (2004) se la plantean.

308-9. Adams (1975) propone la explicación probabilística de las condiciones en que es justificable dar por válidos los condicionales. Entre quienes han propuesto recientemente una visión similar están Evans y Over (2004) y sus colegas, véase, p. ej., Stevenson y Over (1995, pp. 617-618). Hadjichristidis et al. (2001) sostienen que sus resultados corroboran su explicación, no la teoría de modelos. Pero, como señalan Schoryens y Schaeken (2004), ellos malinterpretan la teoría de modelos considerando que implica que quienes razonan construyen modelos completos y no modelos mentales del condicional. Evans et al. (2003) no comprobaron su explicación de una forma directa, porque los participantes en sus experimentos no juzgaron la probabilidad de un condicional y la correspondiente probabilidad condicional. Oberauer y Wilhelm (2003) sí provocaron ambas valoraciones, pero en uno de sus estudios menos de la cuarta parte de los participantes dio las mismas estimaciones para ambas probabilidades. Girotto y Johnson-Laird (2004) no consiguieron encontrar una correlación fiable entre las opiniones de sus sujetos sobre la probabilidad de los condicionales y la probabilidad de la cláusula entonces dada la cláusula si. También variaron en gran medida las estimaciones de los participantes sobre esta última probabilidad. Oaksford y Chater (1996) han extendido su método probabilístico al razonamiento condicional. Mis recelos respecto a este método siguen siendo los mismos (véase el Capítulo 16).

309. Ormerod y sus colegas muestran que los participantes pueden inferir condicionales a partir de disyunciones, y disyunciones a partir de condicionales (véanse Ormerod et al., 1993, y Richardson y Ormerod, 1997). Por ejemplo, dado el condicional:

Si no se van a destinar fondos para la próxima petición de subvención de Phil entonces él estará decepcionado.

podemos inferir la disyunción:

Se van a destinar fondos para la próxima petición de subvención de Phil o él estará decepcionado.

Pero los resultados corroboran una predicción de la teoría de modelos: es más fácil hacer la inferencia conversa partiendo de una disyunción para llegar a un condicional. También confirman la afirmación de Ormerod –plasmada ahora en la teoría de modelos– de que preferimos trabajar con modelos tan parsimoniosos como sea posible. En estudios sin publicar, Sonja Geiger ha encontrado que los participantes que llevan a cabo esta tarea de parafraseo antes de estimar la probabilidad de un condicional tienden a tener en cuenta más posibilidades en sus estimaciones.

309. Girotto et al. (1997) muestran que la dificultad de la inferencia de Chernobyl depende en parte del orden de las premisas. Como también predice la teoría, la inferencia es más fácil con condicionales de la forma: Si y solo si A entonces B, los cuales tienen solo dos modelos completos, que con condicionales de la forma: Si A entonces B, que tienen tres modelos completos en su interpretación lógica (Johnson-Laird et al. 1992).

309. Braine y O'Brien (1991) sostienen que la cláusula si de un condicional implica de forma válida la correspondiente cláusula entonces. Barwise (1986) propone un principio más amplio: la situación descrita en la cláusula si en un condicional restringe la situación descrita en la cláusula entonces.

309. Algunos expertos no han llegado a comprender el hecho de que la teoría de modelos implica que los condicionales pueden tener un número indeterminado de interpretaciones (con todo respeto a Evans y Over, 2004).

Capítulo 22: Causas y obligaciones

311. La cita de Hume aparece en An Enquiry Concerning Human Understanding (1748/1988). Salsburg escribe: “No existe, en realidad, nada similar a causa y efecto. Es una quimera popular, una noción vaga que no resistirá los embates de la pura razón. Contiene un inconsistente conjunto de ideas contradictorias y tiene escaso o nulo valor en el discurso científico” (véase Salsburg, 2001, pp. 185-6; véase también Russell, 1912-13).

311. La teoría de la causación que aparece en este capítulo se publicó por primera vez en Johnson-Laird (1999). La corroboración empírica de la teoría se debe a Yevgeniya Goldvag (ahora Goldvarg- Steingold), véase Goldvarg y Johnson-Laird (2001). Kuhnmünch y Beller (2005) han criticado justamente uno de nuestros experimentos, pero Johnson-Laird et al. (2006a) les han replicado.

312. La idea de la intervención para iniciar una cadena causal se destaca en un análisis de la causación en términos de redes Bayesianas (véase, p. ej., Spirtes et al., 1993, y Pearl, 2000).

312. Sloman (2005) defiende la importancia psicológica de la intervención, y Sloman y Lagnado (2005) presentan pruebas corroborando esta afirmación.

313. El análisis de la causación de Hume se cita a partir de An Enquiry Concernig Human Understanding (1748/1988, p. 115).

314. En relación con las pruebas de que los participantes consideran que las causas preceden a los efectos, véanse Tversky y Kahneman (1980), y Bullock et al. (1982).

314. Debo agradecer el ejemplo del dentífrico a Paolo Legrenzi. Kant (1787/1934) utilizó el ejemplo de una bola de billar que causa un hoyo simultáneo en un cojín (véase también Taylor, 1966).

314. Michotte (1963) demuestra las opiniones perceptivas sobre la causación cuando un objeto golpea a otro. Puede haber restricciones innatas en las percepciones causales (véase, p. ej., Leslie, 1984).

314. Geminiani et al. (1996), haciéndose eco de Hume (1739/1978), sostienen que las concepciones de los adultos sobre la causación física dependen del contacto físico. De forma similar, Lewis (1986) muestra que bastan tres

principios para un programa de ordenador que construya modelos de un sistema físico: todo suceso tiene una causa (!), las causas preceden a sus efectos, y es probable que la causa de cualquier cambio que experimente un objeto sea una acción sobre él. Miller y Johnson-Laird (1976) dan argumentos en favor de una noción de la causación más abstracta que la que se plasma en la percepción visual.

314. Hume (1748/1988) sostiene que una relación causal significa solo una conjunción constante de causa y efecto, mientras que Kant (1787/1934) defiende una conexión necesaria que él consideraba un componente de una concepción innata de la causalidad. En la vida real, cuando afirmamos A causará B, queremos decir que dado A, se debe producir B (véase Harré y Madden, 1975, y los resultados de los que informo posteriormente en el capítulo).

315. En un famoso análisis, Grife (1975) se refiere a inferencias basadas en las convenciones de que quienes hablan dicen la verdad y dan información como "implicaturas". El análisis de los conceptos legales de causación de Hart y Honoré (1985), el análisis filosófico de un "campo causal" de Mackie (1980) y el concepto psicológico de un "conjunto focal" de sucesos de Cheng y Novick (1991) recurren a la idea de que la interpretación causal depende del conocimiento de la situación.

315-6. Mill (1874) niega que exista ninguna distinción de significado entre causas y condiciones posibilitadoras. Hart y Honoré (1985) establecen que la causa es el estado no habitual y la condición posibilitadora es el estado habitual. Cheg y Novick (1991) establecen que la causa es inconstante y la condición posibilitadora es constante. De acuerdo con otros autores, la causa viola una norma asumida por defecto, mientras que la condición posibilitadora, no (véase, p. ej., Kahneman y Miller, 1986; Einhorn y Hogarth, 1986). Y, hay todavía otro grupo de teóricos de acuerdo con los cuales la causa es el factor que es relevante en cualquier conversación explicativa: describimos la causa, no el posibilitador (Mackie, 1980; Turnbull y Slugoski, 1988; Hilton y Erb, 1996).

316. Goldvarg y Johnson-Laird (2001) informan de un estudio en el que los participantes identificaban causas y posibilitadores en diversas situaciones futuras. En un estudio anterior, Cheng y Novick (1991) obtuvieron un resultado similar, pero en sus situaciones las condiciones posibilitadoras eran constantes –se producían en todas las posibilidades– y solo las causas eran inconstantes. Goldvarg y Johnson- Laird informan también del estudio de las inferencias a

partir de premisas causales y posibilitadoras, y de las inferencias a partir de pares de afirmaciones causales.

317. Reichenbach (1956) propone un análisis probabilístico de la causación (véase también von Mises, 1957; Suppes, 1970, 1984; Salmon, 1980). Señala que una causa puede “separar”, es decir, hacer irrelevantes otros sucesos que elevan la probabilidad de su efecto. Por consiguiente, si la probabilidad del efecto, dados la causa y otro suceso, iguala a la probabilidad del efecto dada la causa, entonces el otro suceso es irrelevante (véase el artículo de Hitchcock sobre el tema en plato.stanford.deu/entries/causation-probabilistic/). Cheng y sus colegas defienden una teoría psicológica similar (p. ej., Cheng y Novick, 1990; véase también Schustack, 1988). Hay explicaciones probabilísticas paralelas para el significado de los condicionales (véase, p. ej., Newstead et al., 1997; Oaksford et al., 2000). “Debido a que las relaciones causales no son ni observables ni deducibles”, escribe Cheng (1997, p. 367), “deben inducirse a partir de sucesos observables”. Asume que observamos la diferencia en las respectivas probabilidades a que se refiere la teoría de Reichenbach. Esta teoría no consigue hacer las predicciones correctas para ciertas inducciones causales y por ello Cheng (1997) propone un modelo más sofisticado (el modelo “Power PC”). Las probabilidades desempeñan un papel en la inducción de relaciones causales (véase Pearl, 1988, 2000; Shafer, 1996; Cheg, 1997; Lober y Shanks, 2000). Y la principal prueba a favor de una semántica probabilística es que la gente induce relaciones causales a partir de datos en los que el antecedente no es ni necesario ni suficiente para provocar el efecto (p. ej., McArthur, 1972; Cheng y Novick, 1990; Cummins et al., 1991). Debilitan las generalizaciones causales en las que saben que se producen excepciones, quizá porque saben que muchas causas solo dan lugar a sus efectos si se dan las necesarias condiciones posibilitadoras y no se dan las condiciones que los impiden. Los participantes son sensibles a estos factores, como muestra Cummins (1995, 1998).

317. Hume (1748/1988) defiende la causación como una conjunción constante. Kant (1787/1934, p. 90) escribe que la causación exige “que algo, A, sea de tal modo, que otro algo, B, le siga necesariamente”. Mill (1874, p. 237) escribe: “El antecedente invariable se denomina la causa; el consecuente invariable, el efecto.”

317-8. La redes bayesianas, que son formas eficaces de representar probabilidades (véase, p. ej., Pearl, 2000), no distinguen entre causas y condiciones posibilitadoras. Queda por ver si, en principio, podrían hacer esta

distinción. Si tienen que captar cómo pensamos sobre las relaciones causales necesitan hacerlo.

318. Hume (1739/1978) sostiene que las personas consideran la casualidad como un caso de causa oculta.

319. Muchos autores recalcan el papel de los poderes o mecanismos en el razonamiento causal (p. ej., Harré y Madden, 1975; Ahn et al., 1995; White, 1995; Cheng, 1997): ”un mecanismo causal es el proceso mediante el cual una causa provoca un efecto” (Koslowski, 1996, p. 6). Otros recalcan un marco general de leyes científicas o principios explicativos (Carnap, 1966). Estos principios, sostienen Hart y Honoré (1985), yacen detrás de toda proposición causal sobre sucesos singulares. El poder causal es la idea de que las causas producen efectos mediante una transmisión de energía o alguna otra propiedad que va de la causa al efecto. “En este sentido”, escriben Harré y Madden (1975, p. 5), “la causación siempre implica un material concreto que produce o genera algo”. Cuando se utiliza un martillo para machacar un plato, por ejemplo, el efecto lo produce un poder causal concreto del martillazo: la energía del martillazo se transfiere al plato, causando que se rompa. La teorías del poder causal por tanto definen la causación en función de las propiedades intrínsecas de los objetos. Hume (1739/1978), sin embargo, rechaza las propiedades intrínsecas que producen efectos, argumentando que la definición de producción no se puede distinguir de la de la propia causación.

319. Shanks y sus colegas han explorado las relaciones entre cómo aprenden los organismos las contingencias entre sucesos y cómo hacen inducciones causales (véanse, p. ej., López et al., 1998; Corlett et al., 2004; Shanks, 2004).

319. En relación con estudios que muestran que el conocimiento de los mecanismos causales inunda la información sobre las frecuencias con que aparecen juntos la causa y el efecto, véanse, p. ej., White (1995) y Koslowski (1996).

319. En relación con la explicación del autismo, véase Dawes (2001, p. 136).

320. Mill (1874) defiende la naturaleza caprichosa de la diferencia entre causas y condiciones posibilitadoras, a las que se refiere como “condiciones”. Hart y Honoré (1985) se hacen eco de Mill en su descripción de la causación en la ley. Greene y Darley (1998) sostienen que el Model Penal Code (American Law

Institute, 1980/1985) se basa en asunciones sobre la causación que discrepan de las de los sujetos que no son abogados. A diferencia del Código, consideramos que acciones que son suficientes para provocar la muerte de una víctima exigen el castigo más severo.

320. En cuanto a una discusión sobre las causas “más cercanas”, véase, p. ej., Hart y Honoré (1985, p. 86).

320-1. Para tener una descripción y una discusión de la lógica deóntica, véanse, p. ej., Hilpinen (1971) y van Fraasen (1973).

321. Manktelow (2004) extrae las relaciones entre causación y obligaciones desde un punto de partida diferente.

321. La teoría de modelos para el razonamiento deóntico se desarrolló en colaboración con Monica Bucciarelli (véase, p. ej., Bucciarelli y Johnson-Laird, 2000, 2005).

321. El imperativo categórico se describe en Kant (1785/1997).

321-2. En relación con estudios sobre la comprensión y utilización de proposiciones deónticas por parte de los niños, véanse, p. ej., Hirst y Weil (1982), Byrnes y Duff (1989), Gralinski y Kopp (1993), Day (1996) y Kalish (1998). Girotto, Light y Colbourn (1988) muestran que chicos de dieciocho años pueden resolver versiones deónticas de la tarea de selección en las que los participantes tienen que elegir solo cuál de las dos cartas volver (la llamada “versión reducida” de la tarea de selección). Cummins (1998) corrobora este resultado, y Harris y Núñez (1996) muestran que incluso los niños de cuatro años pueden hacer frente a la tarea. Sin embargo, cualquier factor que haga destacar los contraejemplos mejora el rendimiento en la tarea de selección reducida –incluso con niños de tres años– (Bucciarelli y Johnson-Laird, 2001). Kohlberg (p. ej., 1986) estudió el desarrollo moral de los niños, e, inspirado por Piaget (1965b), postula una serie de estadios diferentes en su captación de la moralidad a medida que adquieren la capacidad de considerar los problemas desde una perspectiva menos egocéntrica. Queda abierta la cuestión de si estos estadios se producen en oposición a un desarrollo general de la moralidad, tanto intuitivo como consciente.

322. La necesidad de que la mente siga el rastro del estatus de los modelos se discute en el Capítulo 3.

322-3. Bucciarelli llevó a cabo el estudio en el que los participantes pensaban en lo que es permisible dados diversos tipos de proposición (Bucciarelli y Johnson-Laird, 2005). Cristina Quelhas y Ruth Byrne también han llevado a cabo experimentos sin publicar sobre el mismo tema. Han mostrado que inferencias de la forma, Si A entonces B, No-B, por tanto, No-A, son más fáciles con condicionales deónticos que con condicionales sobre hechos. Sostienen que este resultado se produce porque, para los condicionales deónticos, los sujetos tienden a construir modelos tanto de lo que es aceptable como de lo que no lo es.

323-5. Bucciarelli llevó también a cabo los experimentos sobre inferencias a partir de pares de premisas deónticas y los estudios sobre inferencias deónticas ilusorias (Bucciarelli y Johnson-Laird, 2005). En los Capítulos 8, 10 y 15 describo otras inferencias ilusorias.

326. Cosmides (1989) es la psicóloga evolucionista que defiende la existencia de un módulo mental especializado en “comprobar el engaño”. Informa de pruebas que, a partir de la tarea de selección, muestran que los problemas referidos a las trampas pueden provocar pautas de respuesta bastante diferentes de las que se producen normalmente en esta tarea. Pero los condicionales que ella utiliza difieren en significado de los condicionales factuales normales –también tienen una “forma lógica” diferente–. Johnson-Laird et al. (1972) informan sobre el primer estudio que utilizó una regla deóntica, el cual se refería a cuánto franqueo poner en distintos tipos de correo. Mostró una enorme mejora en el rendimiento, y hay ya una amplia bibliografía sobre versiones deónticas de la tarea de selección (si se quiere una revisión, véase Manketow, 1999). Los resultados de Carlisle y Shafir (2005) ponen en duda un módulo para el razonamiento sobre el engaño. Perham y Oaksford (2005) muestran que la teoría de decisiones proporciona una explicación del rendimiento en una versión deóntica de la tarea de selección mejor que la que da la teoría evolucionista, cuando la afirmación general se refería a un peligro incluido en una situación futura sobre la posibilidad de engaño, p. ej., “Si hay una herida que sangra entonces el jugador debe dejar el campo” (cf. Fiddick, 2004). También dan argumentos en contra de otras teorías sobre el rendimiento, incluida una versión anterior de la teoría de modelos.

326. En Osherson (1974-76), Rips (1994) y Braine y O’Brien (1998) se presentan descripciones del razonamiento deóntico basadas en reglas formales.

327. Los comentarios de Lord Halsbury pertenecen a Quinn v. Leathem (1901).

Debo agradecer las diversas paradojas de la justicia a Suber (1990).

328. Los dilemas deónticos, como el que se refiere a la operación del niño, son un ingrediente básico tanto de la filosofía moral (p. ej., Harman, 1977; Foot, 2002) como de la psicología moral (p.ej., Kohlberg, 1986; Nichols, 2002).

329. En relación con una explicación sobre Sidney Smith, véase Pearson (1984). La discusión imaginaria entre Buddha y Nietzsche aparece en Russell (1946, pp. 768-773).

329. Osofsky et al. (2005) estudiaron a oficiales de prisiones que llevaban a cabo ejecuciones y sus técnicas para la desconexión moral.

329. Greene et al. (2001) informan del estudio sobre el dilema del tranvía. Utilizaron la resonancia magnética funcional (RMf) mientras los participantes tomaban sus decisiones sobre una serie de dilemas. Aquellos en los que tenían que actuar de una manera directa y personal activaban regiones del cerebro que intervienen en emociones complejas (circunvolución frontal media, circunvolución cingulada posterior y surco temporal superior), mientras que aquellos en los que actuaban de una forma más impersonal activaban regiones que intervienen en la memoria de trabajo (áreas parietal y prefrontal dorsolateral), de manera muy similar a los problemas que no tienen nada que ver con juicios morales. Nichols (2002) también discute el papel de las emociones en los juicios morales.

329. El utilitarismo se debe a Bentham (1789/1996) y Mill (1863/1998). Tanto Baron (2000, capítulo 16) como Sunstein (2005) sugieren que se puede considerar como una explicación normativa de la moralidad, pero que los heurísticos morales –intuiciones basadas en razonamiento inconsciente, dicho con mis palabras– determinan con frecuencia nuestras decisiones, llevándonos algunas veces al error. Por ejemplo, preferimos hacer daño por omisión antes que por acción incluso si las consecuencias utilitaristas van en sentido contrario. ¿Deberíamos vacunar a nuestro hijo contra la tosferina o no? Hay quien prefiere no vacunar a su hijo, incluso siendo menor el riesgo de muerte por la vacuna que por la tosferina (véanse Baron y Ritov, 1994; Royzman y Baron, 2002).

329. Hare (1981) discute los dilemas para los que no hay respuestas correctas o incorrectas.

Capítulo 23: Creencias, herejías y cambios de opinión

331. Gosse describe su idolatría en sus memorias, Father and Son (Gosse, 1923); respecto al libro de su padre, véase Gosse (1998).

331. El estudio sobre el culto del juicio final se describe en When Prophecy Failed (Festinger et al., 1956). Los resultados son controvertidos (véase Bainbridge, 1997).

331. Russell (1957) distingue entre conocimiento directo y conocimiento por descripción. Un filósofo anterior, Reid (1846), se refería al "principio original" del testimonio: tendemos a creer aquellas proposiciones que nos cuenta la gente.

331-2. La versión de la Biblia del Rey Jaime traduce la Epístola a los Hebreos de San Pablo, 11.1: "Ahora bien la fe es la sustancia [hypostasis] de las cosas que esperamos, la prueba [elenchon] de las cosas que no se ven." Humphrey (1996) en su obra Leaps of Faith examina las bases psicológicas de las creencias en lo sobrenatural y ESP (percepción extrasensorial).

332. Un cristiano primitivo que se alegraba de la imposibilidad de sus creencias era Tertuliano. Escribió: Et sepultus resurrexit; certum est, quia impossibile. (Y Él fue sepultado, y se levantó de nuevo: es cierto, porque es imposible; véase 5.4 de Tertuliano, 1956). "Credo quia impossible" (Creo porque es imposible) se atribuye también a San Agustín (véanse sus Confesiones, 1992, VI, 5, 7).

332. La cita sobre la fe es de James (1956, p. 9).

332. En relación con los datos sobre cómo los iguales pueden presionar a los participantes para aceptar proposiciones contrarias a los datos de sus sentidos, véanse, p. ej., Crutchfield (1955) y Asch (1956). En cuanto a la utilización de amenazas de un castigo terrible en el entrenamiento de terroristas, véase Stern (2003), que también es la base para mi manifiesto terrorista multiuso. Ella estudió una variedad de grupos terroristas, incluidos Hammas (musulmán), Kach and Kahane Chai (judío) y the Covenant, the Sword and the Arm of the Lord (cristiano).

332. El estatus de las opiniones sobre Jesús es el siguiente: Jesús era dos personas es la herejía del nestorianismo; Jesús tenía dos naturalezas, una humana

y una divina, es la visión ortodoxa de la iglesia católica; Jesús tenía una única naturaleza divina es la herejía del monofisicismo, y Jesús tenía una única naturaleza humana y era un agente de Dios es la herejía del arrianismo.

333. Para conocer una terrorífica descripción de la Cruzada contra los Cátaros, véase O'Shea (2000).

333. Los Hechos de los Apóstoles dan tres versiones de la conversión de San Pablo (ix, 1-19; xxii, 3-21; xxvi, 9-23).

334. Muchos estudios psicológicos demuestran la perseverancia de los estereotipos sociales frente a pruebas conflictivas (véase, p. ej., Ross y Lepper, 1980; Lepper et al., 1986; Rehder y Hastie, 1996).

334. La versión de Parks sobre la "residenza" aparece en Powers (1997).

335. Oakhill y Johnson-Laird (1985) y Oakhill et al. (1989, 1990) informan de los experimentos sobre los efectos de las creencias en la búsqueda de contraejemplos, incluida la inferencia respecto a los franceses que están en el restaurante. Las creencias afectan a la evaluación de conclusiones dadas (véase Evans et al., 1983). También pueden afectar en primer lugar a la dificultad para construir un modelo (véase Cherubini et al., 1998; Klauer et al., 2000). En general, tienen un efecto mayor sobre las inferencias no válidas que sobre las inferencias válidas.

336. La abducción se describe en el Capítulo 14.

336. Sutherland (1996) es un compendio de las inconsistencias en obras literarias. Explica que la inconsistencia sobre los árboles y Adán se produce entre Génesis (1: 11-12 y 2: 8-9). La discusión de las inconsistencias, accidentales y deliberadas, de Kermode, aparece en su ensayo, "Forgetting" (Kermode, 2003). Los psicólogos han demostrado la dificultad de detectar inconsistencias en los textos (véase, p. ej., Black et al., 1986; Otero y Kintsch, 1992).

337. Russell (1967) cuenta la anécdota sobre el hombre que tenía dudas sobre la fecundidad de las contradicciones.

338. Klein (1998) da argumentos en favor de la separación de las creencias en grupos independientes. Sperber (1996) discute las inconsistencias evidentes en

las creencias culturales.

338. La cita de Whitman pertenece a Hojas de Hierba (véase Whitman, 1977).

339-41. Maria Sonino Legrenzi, Paolo Legrenzi y Vittorio Girotto llevaron a cabo los experimentos sobre la detección y resolución de inconsistencias (véase, p. ej., Jhonson-Laird et al., 2004). Se inspiraron en el día en que el coche no podía arrancar en la pequeña ciudad de Vauvenargues en Provenza. Johnson-Laird et al. (2000) informan del primer experimento sobre la evaluación de la consistencia. Legrenzi et al. (2003) informan sobre el experimento en el que los participantes describían posibilidades. En Johnson-Laird et al. (2000) se informa del primer experimento sobre ilusiones de consistencia, y en Legrenzi et al. (2003), del segundo. En Johnson-Laird et al. (2004) aparece una visión general de estos resultados.

341-2. Los sistemas de razonamiento no monotónicos se han desarrollado en inteligencia artificial. Han resultado alejados del razonamiento humano (para una revisión, véanse, p. ej., Ginsberg, 1987; Brewka et al., 1997). En algunos sistemas, una proposición como que las cerillas frotadas adecuadamente se encienden se considera una idealización por defecto (p. ej., Reiter, 1980). Por tanto, la conclusión de que la cerilla se enciende puede retirarse si hay pruebas en contra, pero sin retirar la asunción por defecto. Otros sistemas, sin embargo, permiten que creencias concretas se revisen a la luz de la inconsistencia (p. ej., Doyle, 1979; De Kleer, 1986). Los filósofos han desarrollado también sistemas para la revisión de las creencias frente a las inconsistencias (p. ej., Harman, 1986; Gärdenfors, 1990; Levi, 1991). Algunos de los problemas sutiles que surgen de un método basado en modelos mentales se discuten en Bonnefon (2004).

342. Harman (1986) distingue dos explicaciones principales para el arraigo de las creencias: las “bases” de una creencia, que son las razones que la apoyan (cf. Doyle, 1979; De Kleer, 1986) y la “coherencia” de una creencia con otras (véase Alchourrón et al., 1985; Gärdenfors, 1988, 1990). Ambos factores pueden ser relevantes para la revisión de las creencias, pero la descripción de Alchourrón et al. no restringe los procesos de revisión en ningún sentido (véase Tennant, 2005). Mi colega Dan Osherson tiene también una prueba sin publicar de la misma importancia. Respecto a la idea del abandono de creencias, Fuhrmann (1997, p. 24) escribe: “... cuando se llega a la elección entre dos candidatas a la eliminación, se debe abandonar la menos arraigada”. Los experimentos han

corroborado este principio (Dieussaert et al., 2000; Politzer y Carles, 2001, Experimento 2).

342-3. Thagard (1989, 1992, 2000) ha implementado un juego de programas que evalúa hipótesis alternativas sobre, pongamos por caso, por qué se extinguieron los dinosaurios, computando su coherencia con las pruebas. El usuario establece una red de nodos que representan las proposiciones de las hipótesis que compiten entre sí, las proposiciones de las pruebas relevantes y la coherencia o incoherencia entre cada par de proposiciones (con diferentes grados de solidez). En otras palabras, la coherencia, a diferencia de la consistencia, se considera como una primitiva no analizada que se presenta con distintos grados de solidez. El programa lleva a cabo el proceso de "satisfacción de restricciones" en el que se rechazan nodos para aumentar la coherencia global del sistema. Siempre y cuando el sistema se estabilice, muestra qué hipótesis son más coherentes con las pruebas. Las proposiciones que describen las pruebas tienen un grado de aceptabilidad propio, que añade una base empírica al sistema, pero pueden ser anuladas por proposiciones más coherentes.

343. En un estudio pionero, Revlis et al. (1971) muestran que cuando un hecho entra en conflicto con la consecuencia de un principio general y con una proposición concreta, los participantes tienden a rechazar la proposición concreta (véase también Revlis y Hayes, 1972; Revlis, 1974; Revlin (sic) et al., 2001). Otros experimentos examinan los efectos del orden de los sucesos sobre las creencias (p. ej., Hogarth y Einhorn, 1992; Schlottmann y Anderson, 1995; Zhang et al., 1997).

343. Se han investigado las conexiones entre la forma de las afirmaciones y su credibilidad. Elio y Pelletier (1997) sugieren que las afirmaciones condicionales pueden ser menos creíbles que las afirmaciones categóricas, y Politzer y Carles (2001) sostienen que cualquier afirmación que contenga una conectiva proposicional puede ser menos creíble que una afirmación categórica. La duda respecto a una disyunción "es producto del hecho bastante trivial de que es más compleja debido a que contiene una conexión y tiene más oportunidades de ser una fuente de error" (p. 224).

343. Uri Hasson llevó a cabo la investigación que muestra que el arraigo no consigue explicar qué creencia abandonan los participantes: pueden abandonar la proposición que calificaron como más creíble dentro de una pareja (véase Hasson y Johnson-Laird, 2006).

344. Girotto et al. (2000) informan de los estudios sobre no correspondencias entre modelos de proposiciones y modelos de hechos. Las no correspondencias entre la sintaxis de las cláusulas en las oraciones pueden tener efectos sobre el razonamiento (véase Evans et al., 1993), pero las no correspondencias, en el presente caso, son entre modelos.

345. La cita es de James (1907, p. 59). Entre los teóricos recientes que ha defendido cambios mínimos en las creencias se incluyen Harman (1986) y Gärdenfors (1988).

345-6. El programa para construir cadenas causales que resuelven las inconsistencias se describe en el Capítulo 22. Johnson-Laird et al. (2004) informan de los experimentos en los que los participantes evaluaban resoluciones causales de las inconsistencias. Los participantes cometían la falacia de la "conjunción" en sus decisiones respecto a las probabilidades de las diversas explicaciones (véase Tversky y Kahneman, 1983, quienes descubrieron la falacia en otro contexto.) Clare Walsh ha llevado a cabo experimentos sin publicar mostrando también que los participantes prefieren explicaciones que van más allá del mínimo.

346. Los modelos completos de una relación de causa-efecto, como por ejemplo "el descargar una pistola causa que no haya balas en la recámara", son los siguientes:

pistola descargada	no balas en la recámara.
pistola no descargada	balas en la recámara.
pistola no descargada	no balas en la recámara.

Facilitan el inferir el efecto (no balas) a partir de la causa (descargada), porque dada la causa el efecto es la única posibilidad. Debería ser más difícil inferir la causa (descargada) a partir del efecto (no balas), porque no hay una única causa en los modelos precedentes, y por tanto permiten otras posibles causas, p. ej., pueden haberse disparado todas las balas. Tversky y Kahneman (1982b) mostraron que los condicionales en los que la cláusula si es la causa del efecto se

consideran más probables que los condicionales en los que la cláusula si es una prueba para la causa.

Capítulo 24: Como resolvemos los problemas

349. En relación con una versión de la perspectiva inspirada por la psicología, véase Kubovy (1986).

349. La opinión de que el insight es un proceso especial se remonta a los psicólogos de la Gestalt alemanes. Ellos sostienen que depende de una "reestructuración" o "recentrado" de la percepción (Köhler, 1947; Wertheimer, 1961). Pero los críticos, tanto iniciales (Bullbrook, 1932) como recientes (p. ej., Perkins, 1981, 2000; Weisberg, 1986), sostienen que no hay nada especial en el insight, y que reestructuración y recentrado son nociones oscuras (Ohlsson, 1984 a,b).

350. Metcalfe y Weibe (1987) son los autores de un estudio que muestra que los participantes pueden valorar lo "templados" que se sienten en sus esfuerzos para resolver problemas. Con problemas utilizados en los estudios sobre insight, no se produjeron aumentos en los índices de temperatura: los participantes se sintieron alejados de la solución hasta el mismo momento en que estalló en su cabeza.

350. Kounios y sus colegas sostienen que la solución a los anagramas depende del insight, y han utilizado escáneres cerebrales para establecer las regiones concretas del cerebro que intervienen en ella (véase, p. ej., Jung-Beeman et al., 2004).

350. De Bono (p. ej., 1992) es quien propone el pensamiento lateral. Henle (1962) sostiene que cuando razonamos no conseguimos atenernos a la lógica. Respecto a la opinión de que el razonamiento es imaginativo, véase el libro de Byrne (2005) sobre el pensamiento contrafactual.

350. La historia corta Pierre Menard, autor del Quijote aparece en Borges (1998).

351. Valéry (1972, p. 324) describe el encuentro entre Degas y Mallarmé.

351. Respecto a una formulación y una defensa de la definición NONCE, véase, p. ej., Johnson-Laird (1993).

351. La versión neodarwinista de la evolución de las especies es una síntesis de la genética moderna con el concepto de selección natural de Darwin. La genética da cuenta de la variación entre los sujetos, y el entorno selecciona aquellas variaciones que son más adaptativas, haciendo posible que los sujetos sobrevivan y se reproduzcan. El proceso generativo no puede guiarse por restricciones adquiridas por los organismos durante su vida (véase Mayr, 1982, p. 537). Es crucial distinguir entre una única aplicación de una estrategia neodarwinista y su uso repetido –un proceso mucho más potente– (Dawkins, 1976). La evolución es así el arquetipo de proceso recursivo: se aplica a sus propios resultados satisfactorios. Es imitado en los programas "genéticos" para encontrar soluciones óptimas a los problemas (p. ej., Holland et al., 1986). Muchos psicólogos proponen teorías neodarwinistas de la creatividad (p. ej., Skinner, 1953; Campbell, 1960; Bateson, 1979; Simonton, 1995).

352. Mayr (1982, p. 354) comenta la teoría de Lamarck. La improvisación musical es inconsciente debido a que no tiene acceso a la memoria de trabajo –una conjetura que se corrobora tanto con un programa de ordenador neolamarckiano para la improvisación musical como por el hecho de que una carga de la memoria de trabajo no parece tener efecto en la calidad o fluidez de las improvisaciones de jazz– (véase Johnson-Laird, 2002b).

353. Duncker (1945) ideó y estudió el problema de la vela.

353. El conductista B. F. Skinner sostiene que los sistemas de enseñanza deberían capacitarnos para aprender sin errores (véase, p. ej., Holland y Skinner, 1961). Un ejemplo contrario de un procedimiento que aprende exclusivamente de los errores es la "propagación del error hacia atrás" (McClelland y Rumelhart, 1986).

354. Katona (1940) parece haber sido el primero en estudiar los problemas de formas. La mayoría de los problemas de formas son sencillos, sin embargo, exigen pensar lo suficiente como para ser materia prima de los libros de problemas (p. ej., Orleans y Orleans, 1983). Louis Lee llevó a cabo los experimentos que investigan los problemas de formas (véase Lee y Johnson-Laird, 2004). Otro tipo de problemas parecido a los problemas de formas son aquellos en los que hay que utilizar jarras de agua de distintos tamaños para

medir una determinada cantidad de agua, pero estos se solucionan mediante una fórmula (véase, p. ej., Luchins y Luchins, 1950).

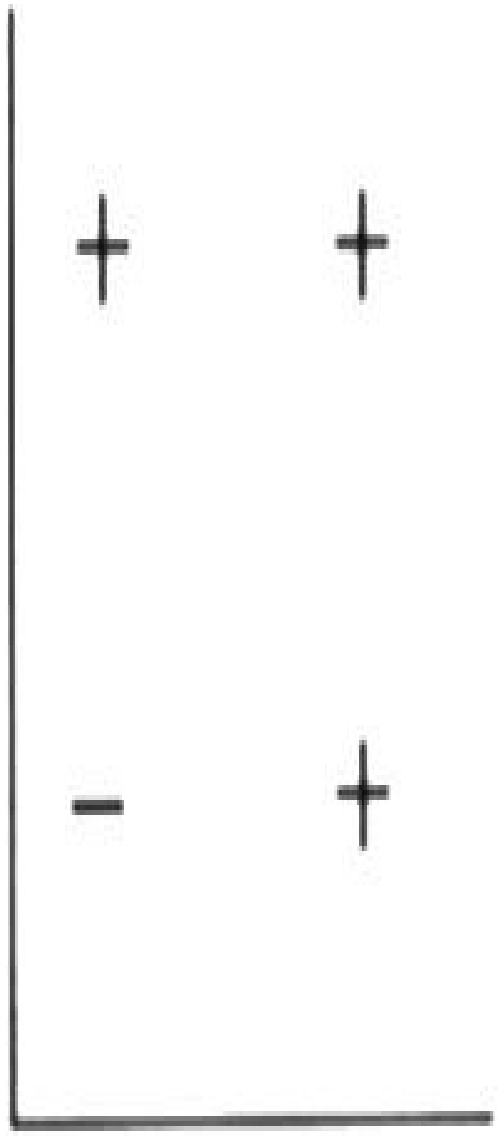
354-5. Louis Lee ideó la explicación respecto a la notoriedad y llevó a cabo los estudios para investigar su efecto, el efecto de la simetría, y el desarrollo de estrategias (Lee y Johnson-Laird, 2004). El desarrollo de estrategias para los problemas de formas sigue el mismo principio general que describo en el capítulo sobre estrategias de razonamiento.

356. La ingeniería inversa se aplica a menudo a programas de ordenador. Se ha visto que es tan habitual que se han hecho intentos para diseñar software que la resista (véase, p. ej., Eilam, 2005).

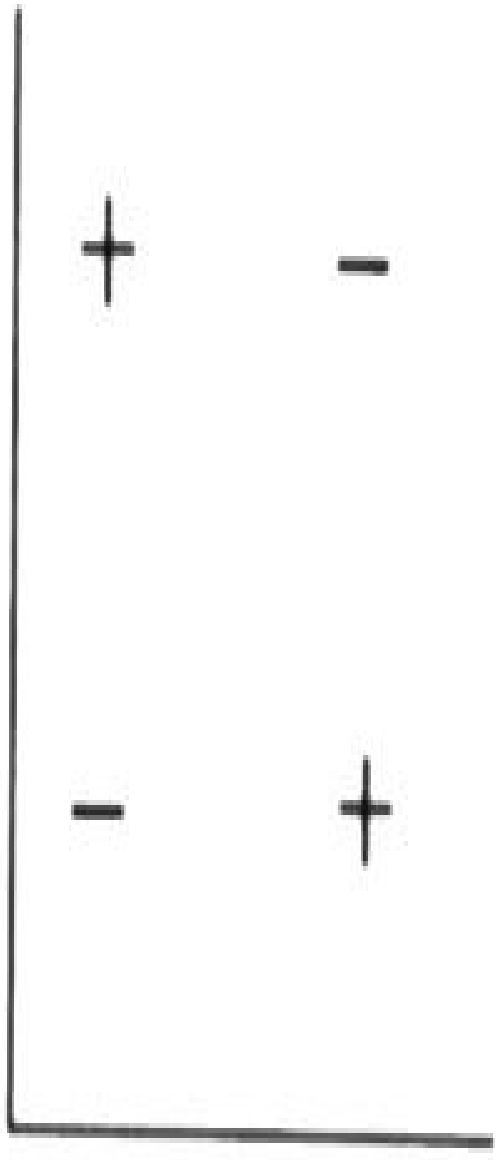
356-7. La teoría de la ingeniería inversa que aparece en este capítulo se presentó por primera vez en Lee y Johnson-Laird (2005).

360-1. Los experimentos, incluido el referido a los grifos de agua, fueron llevados a cabo por Louis Lee (Lee y Johnson-Laird, 2005).

361-2. La noción de independencia, tal como la uso en este capítulo, se relaciona con el hecho de que un concepto se pueda aprender utilizando una herramienta que sea “lineal”. Los programas de ordenador que implementan un procedimiento lineal para aprender se llaman “perceptrones” (Rosenblatt, 1958, 1961). Consideremos la representación espacial de un problema “o” incluyente:



Podemos dibujar una línea recta entre el conjunto de casos en la que la bombilla está encendida (los signos más en la figura) y el caso en que está apagada (el signo negativo). Un perceptrón es una herramienta que aprende a dibujar esa línea recta: es una herramienta “lineal”. Ahora, consideremos la representación espacial de una disyunción excluyente, es decir, un problema “o bien”:



Es imposible dibujar una única línea recta que separe los casos positivos y negativos. Ningún perceptrón puede representar el problema “o bien” (Minsky y Papert, 1969). Los programas que pueden aprender disyunciones excluyentes son por tanto “no lineales” (véase, p. ej., McLelland y Rumehart, 1986; Vapnik, 1998). A diferencia de esta distinción, la noción de independencia puede tener distintos grados (véase el texto de este capítulo).

362. El análisis medios-fines se describe en Polya (1973), y respalda el trabajo fundamental de Newell y Simon (1972) sobre la solución de problemas por parte de los humanos.

362-3. Weisberg (1996) clasifica los problemas de “insight” de una forma preteórica. Las recientes teorías sobre el insight incluyen la opinión de que depende de la superación de un enfoque erróneo de un problema (véase, p. ej., Wisberg, 1993; Isaak y Just, 1995; Knoblich et al., 1999). Un período de “incubación” en el que pensamos en otros asuntos permite que los indicios que conducen al error se hagan menos accesibles con la consiguiente mayor opción de que recobremos los indicios correctos (Smith y Blankenship, 1991). Realmente perseveramos (véase, p. ej., Jansson y Smith, 1991). Sin embargo, el problema con este método es su asunción de que conocemos los métodos correctos, aunque no podamos tener acceso a ellos. La afirmación puede ser correcta para algunos problemas (véase Keane, 1989), pero es falsa para otros. Otra visión del insight se basa en el concepto de Newell y Simon (1972) de un “espacio del problema”, es decir, el espacio abstracto definido por todas las posibles secuencias de las operaciones mentales relevantes para un problema. Quien resuelve el problema necesita buscar este espacio para una secuencia de operaciones que lleva desde el estado inicial del problema a su solución. Kaplan y Simon (1990) sostienen que el insight depende de un cambio desde una búsqueda en el espacio del problema a una búsqueda en el espacio del metanivel de posibles espacios del problema para tener una nueva representación del problema (véase también Korf, 1980). En un caso de insight, esta nueva representación da lugar a la solución del problema. Pero, como Kaplan y Simon señalan (p. 381), “El espacio de posibles espacios del problema es extremadamente indefinido, en realidad, ‘infinito’” (p. 381).

363-4. Otra opinión más sobre el insight se debe a Tom Ormerod y sus colegas: cuando tratamos de resolver un problema, seleccionamos pasos para maximizar el avance hacia el objetivo pero minimizar la expansión del espacio del problema (véase Chronicle et al., 2001; MacGregor et al., 2001; Ormerod et al., 2002). Relajamos esta restricción solo si tenemos que hacerlo, pero de esa forma podemos descubrir un nuevo tipo de paso. Esta teoría postula que podemos evaluar si un paso putativo supone un avance hacia un objetivo.

Capítulo 25: Bicicletas voladoras: cómo inventaron el aeroplano los hermanos Wright

365. El epígrafe de la Parte VIII pertenece a Auden (1977, p. 404).

367. Existe una amplia bibliografía sobre los hermanos Wright. He considerado indispensables sus propios escritos, p. ej., Kelly (1996), Wright (1988) y la magistral colección de McFarland (2001) respecto a su correspondencia y artículos, la biografía definitiva de Crouch (1989), la monografía técnica de Jakab (1990) y la reciente descripción de la carrera por volar de Tobin (2003). En cuanto a una versión independiente de sus logros desde el punto de vista de la ciencia cognitiva, véase Bradshaw (1992).

367. El 15 de octubre de 1906 Wilbur escribe a otro pionero de la aviación, Octave Chanute (Kelly, 1996, p. 183): “Si las ruedas del tiempo pudieran retroceder seis años atrás, no es en absoluto probable que hiciéramos de nuevo lo que hemos hecho. Lo único que me impresiona como destacable es el poco tiempo en que se hizo nuestro trabajo.”

368. Tobin (2003) describe los esfuerzos sin éxito de Langley y su coste.

368. El comentario de Kelvin aparece en una carta a la mujer de Alexander Graham Bell en la que se refiere a los esfuerzos del propio Bell respecto a la aviación (citado en Tobin, 2003, p. 251).

369. Jakab (1990, p. 26) incluye la cita de Maxim sobre los motores para aviones.

369. La afirmación de Wilbur sobre el conocimiento y la habilidad es de McFarland (2001, p. 15).

369. El 20 de abril de 1903 Wilbur escribe en una carta a un colega, George A. Spratt: “Es una característica de toda nuestra familia el ser capaces de ver los puntos débiles de cualquier cosa, pero no es siempre una cualidad deseable porque nos hace demasiado conservadores para ser hombres de negocios con éxito, y limita nuestras amistades a un círculo muy limitado” (McFarland, 2001, Vol. 1, p. 306).

370. En cuanto a las relaciones de Wilbur con Montgomery, véase, p. ej., Crouch (1989).

371. Describo las principales teorías respecto a la analogía en el Capítulo 14, y discuto allí la analogía de Wilbur entre la caja de la cámara y las alas de un biplano. Dunbar ha estudiado el pensamiento de los científicos de cuatro laboratorios líderes en biología molecular (p. ej., Dunbar, 1995). Cuando fallaba un experimento, los científicos buscaban razones en analogías con otros experimentos. En todas las reuniones a las que asistió Dunbar extrajeron estas analogías. También utilizaban analogías entre un dominio y otro pero solo cuando estaban pensando sobre una teoría o planificando una serie de experimentos. Gick y Holyoak (1983) muestran que la gente no siempre entiende el alcance de una analogía.

371-2. En cuanto al comentario de Wilbur respecto a los buitres, véase McFarland (2001, p. 15 et seq.).

372. La descripción de la analogía de la caja de la cámara hecha por Wilbur aparece en una carta a Chanute (véase McFarland, 2001, Vol. 1, p. 18).

372. El 8 de junio de 1907 Wilbur escribe a su hermana desde París (Kelly, 1996, p. 212): “Mi imaginación dibuja las cosas más vívidamente que mis ojos.” Crouch (1989, p. 221) al describir a los hermanos se refiere a “su genialidad para visualizar soluciones mecánicas a problemas teóricos”. Otros autores entre los que se incluyen Jakab (1990) y Ferguson (1992) hacen comentarios similares. Craig et al. (2002) discute el uso de la simulación perceptiva en la analogía. Gooding (2005), el decano de los científicos con beca Faraday, plantea reglas de inferencia visuales para el pensamiento científico, pero sus ejemplos parecen depender de la manipulación de modelos, p. ej., la tarea de reconstruir la forma

tridimensional de un ser vivo a partir de su fósil plano. Gould (1989, p. 92) describe al investigador como “dando vueltas a la maldita cosa en mi cabeza” desde su posición en un dibujo y el ángulo diferente de otro dibujo (cf. mi discusión de la rotación mental de imágenes en el Capítulo 2).

375. El comentario de Wilbur sobre el centro de presión del aire aparece en la transcripción de su primera charla profesional, que dio en la Sociedad Occidental de Ingenieros [Western Society of Engineers] en Chicado en septiembre de 1901 (véase p. 101, de “Some Aeronautical Experiments” , en McFarland, 2001, Vol. 1, pp. 99-118).

376. “Ni en cien años...” se cita en Crouch (1989, p. 213).

377. Los hermanos no fueron los primeros en construir un túnel de viento, pero fueron los primeros en utilizar un túnel de viento para hacer medidas exactas útiles para el diseño de una aeronave (véase, p. ej., McFarland, 2001, Vol. 1, p. 548, f.n. 5).

378-9. Los comentarios sobre el diseño de hélices de los Wright aparecen en su artículo “The Wright Brothers’ Aeroplane” (Wright, 1988). Combs (1979, p. 182) sostiene que su teoría de las hélices les situó diez o veinte años por delante de sus rivales, y es cierto que los rivales de los Wright continuaron utilizando el ensayo y error en sus diseños de hélices durante al menos diez años. En cuanto a la diferencia con las hélices de los barcos, véase McFarland (2001, Vol. 1, p. 596).

379. Los cálculos maratonianos se muestran en McFarland (2001, Vol. 1, pp. 49-50).

381. El breve obituario de su padre para Wilbur no ha sido superado: “Esta mañana a las 3:15 falleció Wilbur, tenía 45 años, 1 mes y 14 días. Una vida corta, llena de consecuencias. Una inteligencia a toda prueba, un carácter imperturbable, una gran independencia y una modestia igual de grande, viendo claramente lo correcto, persiguiéndolo de manera constante, vivió y murió (del diario de su padre, citado en Crouch, 1989, p. 449).

382. “No es asombroso...”, de una carta que Orville escribió a George A. Spratt, el 7 de junio de 1903 (McFarland, 2001, Vol. 1, pp. 310-315).

Capítulo 26: Desvelar un enigma

383. Pundits distingue entre códigos, p. ej., el código Morse, y cifrados, p. ej., el resultado de la máquina Enigma. Yo utilizo los dos términos de manera intercambiable.

383. La bibliografía sobre los cifrados es amplia. He encontrado útiles estos libros: Hinsley (1975), la historia oficial del Gobierno; Hinsley y Stripp (1993), un conjunto de recuerdos de la gente de Bletchley Park; Kahn (1996), la principal historia de los códigos y los cifrados, y Budiansky (2000), una descripción del descifrado de códigos durante la Segunda Guerra Mundial. Una búsqueda con Google en la Web da como resultado varios sitios que permiten operar con una máquina Enigma simulada.

383. Marian Rejewski y un grupo de criptoanalistas polacos descifraron el código Enigma en 1933 (véase Kozaczuk y Straszak, 2004). Actuaban de enlace con los franceses, y el 25 de julio de 1939 en una reunión conjunta en Pyry, cerca de Varsovia, con expertos franceses e ingleses entre los que estaba Dillwyn Knox (el futuro jefe de Welchman), los polacos con gran generosidad entregaron duplicados de máquinas Enigma, diagramas de las conexiones de los rotores y su método para descifrar el código. Knox había pensado que Enigma era indescifrable. No es de extrañar que estuviera tan furioso porque Welchman, trabajando por su cuenta, descubriera cómo descifrar el código. Las contribuciones de Turing se describen en Hodges (2000).

383. Mi explicación de los métodos de Welchman se basa en el Capítulo 4 de su libro The Hut Six Story (Welchman, 1982). Lo que ha complicado la tarea de los historiadores es que el trabajo en Bletchley permaneció “clasificado” por el gobierno británico durante muchos años. El primer libro en inglés sobre ese trabajo no se publicó hasta 1974, ¡y defendía que los británicos fueron los primeros en descifrar el código Enigma!

384. Edgar Allan Poe invitó a los lectores de un semanario de Filadelfia a enviar mensajes cifrados, y descifró todos ellos utilizando un análisis de frecuencias. En su historia corta El escarabajo de oro (Poe, 1982), el protagonista utiliza ese método para descifrar la descripción que hace el Capitán Kidd de la situación del tesoro enterrado. Conan Doyle también hace a Sherlock Holmes utilizar el

mismo método en La aventura de los muñecos danzantes (Conan Doyle, 1930).

Capítulo 27: Sobre el modo de transmisión del cólera

397. Mi interés por Snow fue suscitado por la maravillosa novela de Keith Oatley (1998) A natural history, cuyo protagonista es un doctor victoriano copiado en parte de Snow. Mi descripción de Snow se basa en los Snowflakes, cinco autores que analizan la carrera y el pensamiento de Snow. Son el historiador Peter Vinten-Johansen, el filósofo doctor en medicina Howard Brody, el epidemiólogo Nigel Planeth, el historiador Stephen Rachman y el epidemiólogo Michael Rip (véase Vinten-Johansen et al., 2003). Sostienen que Snow se basó en el método hipotético-deductivo, que utilizó la inducción para establecer ciertos hechos, y que dedujo consecuencias empíricas de su teoría. Sus investigaciones para comprobar estas consecuencias operaban en diferentes niveles de análisis que iban desde la sociedad en general a interacciones químicas. Califican algunas de las inferencias de Snow como deducciones y otras como inducciones. La teoría de modelos sobre el razonamiento, espero, encaja con su explicación.

398. Porter (1997) es una espléndida historia de la medicina, que incluye sus numerosas modas pasajeras.

398. La primera explicación de Snow sobre el cólera fue On the Mode of Communication of Cholera en agosto 1849a, y dos meses más tarde publicó en una revista un artículo sobre la patología y transmisión de la enfermedad (Snow, 1949b; véase también Frost, 1965).

398. Popper (1959) describe su filosofía de la ciencia.

399. Howson y Urbach (1993, p. 1) sostienen que “el razonamiento científico y... gran parte del razonamiento cotidiano se lleva a cabo en términos probabilísticos”.

399. Herschel (1987) describe su método hipotético-deductivo en un libro que se ganó la aprobación de Darwin. La versión de Herschel fue criticada por filósofos de la ciencia posteriores, quienes tenían un punto de partida más inductivista,

como Whewell (1967) y Mill (1874). Whewell se interpreta con frecuencia de forma errónea como una anticipación de Popper.

399. En relación con una explicación sobre el papel de los modelos en el razonamiento científico, véase Hesse (1966) y Wise (1979), que muestran su papel en el desarrollo de las ecuaciones de campo de Maxwell. Los científicos cognitivos defienden su papel, junto al de la imagen visual, en el desarrollo de las teorías científicas (Miller, 1984; Giere, 1988; Gorman, 1992; Thagard, 1999; y las colecciones de artículos en Giere, 1992; Magnani et al., 1999; Magnani y Nersessian, 2002; Gorman et al., 2005). Simon y sus colegas describen el papel de la inducción en el desarrollo y comprobación de las leyes científicas (véase, p. ej., Langley et al., 1987; Kulkarni y Simon, 1988; Klahr y Simon, 1999).

399. En los Capítulos 5 y 13 discuto el papel del heurístico de Hume, “efectos similares dependen de causas similares”.

399. Porter (1997, capítulo 13) describe las reformas sanitarias y la medicina pública en Inglaterra durante el siglo diecinueve.

400. La cita sobre la ley inversa de los gases es de Snow (1855) y está citada por Vinten-Johansen et al. (2003, p. 345).

400. Vinten-Johansen et al. (2003) discuten los factores subyacentes en la conversión de Snow desde los efluvios nocivos al agua contaminada como forma de transmisión de la enfermedad. Comentan la reflexión de Snow sobre Albion Terrace: “Snow pensaba en términos espaciales y topográficos, aunque él mismo no hiciera diagramas” (p. 208).

406. Respecto a que el mérito de descubrir el Vibrio cholerae lo obtuvo Koch, véase p. ej., Porter (1997) y Halliday (1999); y respecto a la verdadera historia, véase Tognotti (2000, p. 244).

406. La “gran fetidez” del Támesis se describe en la descripción de Halliday (1999) sobre el prodigioso proyecto de ingeniería que representó el trabajo de Sir Joseph Bazalgette para hacer del Támesis el río metropolitano más limpio del mundo –un proyecto que el Parlamento se vio obligado a solicitar como consecuencia de la gran fetidez.

Capítulo 28: Cómo razonamos

407. Las inferencias en el bar se describen detalladamente en el Capítulo 13.

407. El razonamiento inconsciente y las emociones se analizan en la Parte II del libro.

409. La modulación de la interpretación de las conectivas y otros términos lógicos se analiza en el Capítulo 12. Respecto a la imposibilidad de captar las propiedades lógicas de términos relacionales, véase el Capítulo 9. Debo agradecer el ejemplo de Alicia y la poción a Steve Sloman.

409. Rips (1994) comenta la naturaleza fortuita de los errores según su teoría.

410. Boole (1854) analiza las leyes del pensamiento.

410. La idea de que el razonamiento depende de muchos procesos diferentes, como una navaja suiza, se debe a los psicólogos evolucionistas.

410-1. Las ideas de Peirce y la incapacidad de los iconos para representar la negación se describen en el Capítulo 2 (en relación con algunas de sus otras limitaciones, véanse los Capítulos 8 y 10).

411-2. Las inferencias ilusorias se producen en el razonamiento con conectivas (Capítulos 8 y 23), cuantificadores (Capítulo 10), probabilidades (Capítulo 15) y términos deónticos (Capítulo 22).

412. Oaksford y Chater (1998, p. 110) argumentan que el papel de los contraejemplos no es ni necesario ni suficiente para la racionalidad. No es necesario porque durante períodos de ciencia “normal” (cf. Kuhn, 1970) los científicos no permiten que los principios teóricos centrales se invaliden mediante contraejemplos. Puede ser; pero yo hablo de la utilización de contraejemplos para invalidar inferencias, no afirmaciones generales. También argumentan que los contraejemplos no son suficientes para la racionalidad porque no sería racional seguir buscando contraejemplos de manera indefinida cuando estamos tratando de llegar a una decisión en tiempo real. Este argumento, sin embargo, se basa en una asunción que la teoría de modelos no hace, a saber, que las raíces de la racionalidad están en una búsqueda completa

de contraejemplos. En realidad, la teoría postula que somos racionales porque captamos la fuerza de los contraejemplos. Hasta puede que los busquemos ocasionalmente, aunque no parece que tengamos ningún procedimiento sistemático, menos aún abarcable, para buscarlos. Es sorprendente que la base de la crítica de Oaksford y Chater sea pensar en contraejemplos, aunque sea para una afirmación más rotunda que la que yo hacía.

413. Ragni (2003) prueba que el razonamiento espacial bidimensional es inabarcable.

413. James (1907, p. 59) argumenta en favor de los cambios mínimos en nuestras creencias.

413-4. Stanovich (1999) informa de muchos estudios que establecen correlaciones entre razonamiento e inteligencia.

414. Simon (1982) sostiene que la capacidad de la memoria de trabajo suponía un importante cuello de botella en el razonamiento. En cuanto a estudios que relacionan la capacidad de la memoria de trabajo con la capacidad para razonar, véase, p. ej., Bara et al. (1995), Barrrouillet y Leças (1999).

414. Las memorias de U. S. Grant, escritas cuando se estaba muriendo, revelan que fue un gran general y un gran escritor (véase Grant, 1990). Wellington no escribió memorias, pero su inferencia sobre la existencia del vado a través del río, que sus guías habían negado, se describe en la biografía de Longford (1969, p. 90).

414. Cito el comentario de Whitman (1977) sobre la contradicción en el Capítulo 23.

414. James Joyce (1930) escribe en The Portrait of the Artist as a Young Man: “... las únicas armas que me permito utilizar... silencio, exilio y astucia”.

414. El Paris Review, editado durante muchos años por el difunto George Plimpton, ha publicado muchas entrevistas con poetas y escritores. No recuerdo que ninguno de ellos se refiriera al papel del razonamiento en su trabajo.

414-5. Auden discute la creación de poesía en el Prólogo de La mano del teñidor (Auden, 1962; véase también Auden, 1973, p. 358 et seq.). Spender (1952) describe cómo escribió un poema. Vendler (2004) describe el pensamiento en la

poesía de Pope, Whitman, Dickinson y Yeats.

415. La cita es de Valéry (1958, p. 58).

Notas al pie

[1] Nota de la tr.: En el original inglés “any” y “some”.

Agradecimientos

Un libro basado en la investigación psicológica debe tanto a todos aquellos que han colaborado en la investigación que el principal papel del autor es encauzar sus esfuerzos. Parece muy poca cosa hacer una lista con sus nombres por orden alfabético. Si pudiera contarle algo más sobre cada uno de ellos sin abusar de su paciencia, lo haría, pero es imposible al final de un libro extenso. Sin embargo, debo señalar de manera singular a varios grupos de amigos con los que he llevado a cabo investigaciones a lo largo de muchos años: Ruth M. J. Byrne, vicerrectora del Trinity College en la Universidad de Dublín; Bruno Bara y Monica Bucciarelli, de la Universidad de Turín; Juan García-Madruga y sus colegas de la Universidad de Educación a Distancia, Madrid; Paolo Legrenzi, Maria Sonino Legrenzi y Vittorio Girotto, de la Universidad de Arquitectura, Venecia; Keith Oatley, profesor emérito en la Universidad de Toronto; y Francesco Mancini, el director de la Escuela de Psicoterapia, Roma, y Amelia Gangemi, de la Universidad de Cagliari, Cerdeña.

A continuación, mi agradecimiento es para las seis personas que leyeron el manuscrito completo del libro en su versión inicial. Geoff Goodwin y Louis Lee añadieron sus propias investigaciones y revisaron todo el libro con lupa. Agradezco también a Louis el que haya preparado unas figuras decentes a partir de los resultados de mi viejo programa de gráficos. Vittorio Girotto me sugirió muchas citas oportunas, especialmente en lo referente a bibliografía sobre psicología animal, y me señaló muchas afirmaciones caprichosas, que amenazó con discutirme en otras publicaciones. Monica Bucciarelli me puso directamente en contacto con aspectos de la bibliografía sobre desarrollo infantil, y me avisó de las dificultades idiomáticas del libro para quienes no fueran angloparlantes. Ruth Byrne se las arregló para combinar la lectura del libro con su puesto de segunda de a bordo en el Trinity College de Dublín. Hizo muchas críticas perspicaces y me animó mucho, excepto cuando escribí “cogniciones” como nombre en plural. Envié el manuscrito a otros cuatro expertos, que estaban demasiado ocupados como para leerlo. No mencionaré sus nombres, pero ellos saben quiénes son, y son responsables de cualquier error que reste en el libro. (¡Es broma!)

Muchas otras personas han realizado una contribución esencial colaborando en la investigación, y quiero darles las gracias: Patricia Barres, Malcolm Bauer, Victoria Bell, Nuria Carriedo, Jean-Paul Caverni, Paolo Cherubini, Jon Cohen, Orlando Espino, Jonathan Evans, Rich Feit, Caren Frosch, Alan Garnham, Sonja Geiger, Genya Goldvarg-Steingold, Geoff Goodwin, Francisco Gutiérrez, Simon Handley, Uri Hasson, Cathy Haught, Georg Jahn, Mark John, Csongor Juhos, Jim Kroger, Jung Min Lee, Louis Lee, Rachel McCloy, Mark John, Markus Knauff, Robert Mackiewicz, Henry Markovits, Marc Marschark, Hans- Georg Neth, Mary Newsome, Jane Oakhill, Tom Ormerod, Stefania Pighin, Cristina Quelhas, Marco Ragni, Sergio Moreno Rios, Carlos Santamaría, Fabien Savary, Walter Schaeken, Walter Schroyens, Victoria Shaw, Vladimir Sloutsky, Patrizia Tabossi, Alessandra Tasso, Isabelle Vadeboncoeur, Jean-Baptiste van der Henst, Clare Walsh, Yingrui Yang y Dan Zook

Estoy agradecido a Deborah Prentice, la jefe del Departamento de Psicología de Princeton, que ha ayudado a hacer que sea un lugar tan agradable para enseñar e investigar. También me ayudó a conseguir un semestre sabático en el otoño de 2004, que me permitió empezar a trabajar en el libro. Varios compañeros de Princeton han sido una gran fuente de consejo y ánimo: Jon Cohen, Ron Comer, John Darley, Susan Fiske, Harry Frankfurt, Jack Gelfand, Sam Glucksberg, Adele Goldberg, Gil Harman, Norman y Leonore Itzkowitz, Danny Kahneman, Virginia Kwan, George Miller, Dan Osherson, Eldar Shafir, Joe Simmons, Alex Todorov, Ann Tresiman y Edwin Williams.

Gracias a muchos otros expertos por su amable consejo, ánimo y franca crítica: Alexandre Banos, Jon Baron, Pierre Barrouillet, Hart Blanton, Luca Bonatti, Melissa Bowerman, Selmer Bringsjord, Jerome Bruner, Nick Chater, Michelle Cowley, Susan Carey, Michel Denis, Ronnie de Sousa, Kristien Dieussaert, Suzanne Egan, Jonathan St. B. T. Evans, Carol Feldman, Wayne Gray, David Green, Tony Greenwald, Justin Halberda, Earl B. Hunt, Ray Jackendoff, W. Jake Jacobs, Jenny Jenkins, Marcia Johnson, Mak Keane, Rick Lewis, Alice McEleney, Ken Manktelow, Gary Marcus, Julie Meehan, Aisling Murray, Ray Nickerson, Ira Noveck, Mike Oaksford, Kieron O’Connor, David Over, Luis Pineda, Thad Polk, Pat Rabbitt, Carol Raye, Lance Rips, Susana Segura, Steve Sloman, Ed Smith, Dan Sperber, Keith Stenning, Len Talmy, Valerie Thompson, Alexander Todorov, Barbara Tversky, André Vadierendonck y Massimo Warglien.

Varias personas me han ayudado en cuanto al aspecto técnico y a la publicación.

Mi agradecimiento a Martin Bum y Angela Butterworth de OUP por facilitar el camino de la publicación. Mi agradecimiento también para Bernie VanUiter y Marion Kowaleski por la ayuda administrativa, a Donna O'Leary y Keisha Craig por su ayuda como secretarias, y a David Berkowitz, Paul Bree, Carol Agans, Richard Reiss, Jason Robinson y Jim Watson por la ayuda técnica. Parte de la investigación en que está basado este libro fue posible gracias a una beca de la National Science Foundation (Beca 0076287) para el estudio de estrategias de razonamiento. El capítulo 25 sobre los hermanos Wright está basado en parte en un artículo (Johnson-Laird, 2005). Agradezco a mi hermano Andy Johnson-Laird –un intrépido piloto– sus consejos sobre el vuelo y la lectura crítica de ese artículo, y a Nancy Nersessian –una intrépida teórica sobre modelos mentales en ciencia– sus críticas a un primer borrador.

Tengo una gran deuda con mi familia cercana. Mi querida esposa, Mo, fue la sexta persona que se leyó el libro completo. Habló en representación del lector medio, y, como resultado, eliminé porciones completas de incomprensible verborrea. Nuestro hijo Ben hizo los dibujos con el DesignWorkshop_Lite para el capítulo 2 y nuestra hija Dorothy nos proporcionó breves sesiones de psicoterapia cuando fueron necesarias. Mi reconocimiento final es para tres mentores intelectuales, lamentablemente ya fallecidos: Peter Wason, Stuart Sutherland y Jonck, es decir, A. R. Jonckheere.

Referencias

ADAMS, E. W. (1975), The Logic of Conditionals: an Application of Probability to Deductive Logic. Dordrecht: Reidel.

AHN, W.; KALISH, C. W.; MEDIN, D. L., y GELMAN, S. A. (1995), «The role of covariation versus mechanism information in causal attribution», Cognition, 54, 299-352.

ALCHOURRÓN, C.; GÄRDENFORS, P., y MAKINSON, D. (1985), «On the logic of theory change: partial meet contraction functions and their associated revisión functions», Journal of Symbolic Logic, 50, 510-530.

ALEMÁN, A. (2005), «Feelings you can’t imagine: towards a cognitive neuroscience of alexithymia», Trends in Cognitive Sciences, 9, 553-555.

ALLOY, L. B., y ABRAMSON, L. Y. (1979), «Judgement of contingency in depressed and non-depressed students: sadder but wiser?», Journal of Experimental Psychology, 108, 441-485.

ALLOY, L. B., y ABRAMSON, L. Y. (1988), «Depressive realism: four theoretical perspectives». En Alloy, L. B. (ed.), Cognitive Processes in Depression. Nueva York: Guilford Press, pp. 223-265.

AMERICAN LAW INSTITUTE (1980/1985), Model Penal Code and Commentaries. Philadelphia, PA: American Law Institute.

ANDERSEN, H. C. (1996), The Snow Queen and Other Fairy Tales. Londres: Bloomsbury. [Ed. cast.: La Reina de las Nieves y otros cuentos, trad. de A. Adell, Madrid, Alianza, 2004.]

ANDERSON, J. R. (1978), «Arguments concerning representations for mental imagery», Psychohgical Review, 85, 249-277.

ANDERSON, J. R. (1990), The Adaptive Character ofThought. Hillsdale, NJ: Erlbaum.

ANDREWS, G.; CREAMER, M.; CRINO, R.; HUNT, C.; LAMPE, L., y PAGE, A. (2003), The Treatment of Anxiety Disorder: Clinician Guides and Patient Manuals, 2ª ed. Cambridge: Cambridge University Press.

ANGLUIN, D. (1978), «Inductive inference of formal languages from positive data», Information and Control, 45, 117-135.

ANNAS, J., y BARNES, J. (1985), The Modes of Scepticism: Ancient Texts and Modern Interpretations. Cambridge: Cambridge University Press.

ANNO, M., y NOZAKI, A. (1984), Anno's Hat Tricks. Londres: Bodley Head.

AREVALO, Z., y BREDESON, J. G. (1978), «A method to simplify a Boolean function into a new minimal sum-of-products for programmable logic arrays», IEEE Transactions on Cornputing, C- 27, 1028-1030.

ARISTÓTELES, Tratados de Lógica (trad. M. Candel Sanmartín). Madrid. Ed. Gredos, 1988.

ASCH, S. E. (1956), «Studies of independence and conformity: a minority of one against a unanimous majority. Psychological Monographs, 70 (nº 416 completo).

ASHLCY, K. (1990), Modeling Legal Arguments: Reasoning with Cuses and Hypotheticals. Cambridge, MA: MIT Press.

AUDEN, W. H. (1962), The Dyer's Hand and Other Essays. Nueva York: Random House. [Ed. cast.: La mano del teñidor y otros ensayos, trad. de A.. Oquendo y M. Lauer, Barcelona, Barral Editores (1974).]

AUDEN, W. H. (1971), A Certain World: a Commonplace Book. Londres: Faber & Faber.

AUDEN, W. H. (1973), Forewords and Afterwords. Nueva York: Random House. [Ed. cast.: Prólogos y Epílogos, trad. de M. Martínez Lage, Barcelona, Ediciones Península (2003).]

AUDEN, W. H. (1976), Mendelson, E. (ed.), Collected Poems. Nueva York: Random House.

AUDEN, W. H. (1977), Mendelson, E. (ed.), The English Auden: Poems, Essays and Dramatic Writings 1927-1939. Londres: Faber & Faber.

AGUSTÍN, SAN (1985), La ciudad de Dios (trad. J. C. Díaz de Beyral), Ediciones Orbis.

AGUSTÍN, SAN (2007), Confesiones (trad. P. A. de Urbina), Ediciones Palabra.

AUSTIN, J. L. (1970), «Truth». En Urmson, J. O., y Warnock, J. G. (eds.), Philosophical Papers of J.L. Austin, 2ª ed. Oxford: Oxford University Press, pp. 117-133.

BACON, F., Ensayos (trad. L. Escolar Bareño), Aguilar. (Publicado originalmente en 1597.)

BACON, F. (2003), Novum organum (trad. C. Litrán), Ediciones Folio. (Publicado originalmente en 1620.)

BADDELEY, A. D. (1986), Working Memory. Oxford: Clarendon Press.

BADDELEY, A. D. (2000), «The episodic buffer: a new component of working memory», Trends in Cognitive Sciences, 4, 417-423.

BAGULEY, T., y PAYNE, S. J. (2000), «Given-new versus new-given? An analysis of reading times for spatial descriptions». En Nualláin, S. Ó. (ed.), Spatial Cognition. Amsterdam: Benjamins, pp. 317-328.

BAILLARGEON, R. (1987), «Young infants' reasoning about the physical and spatial characteristics of hidden objects», Cognitive Development, 2, 178-200.

BAILLARGEON, R., y HANKO-SUMMES, S. (1990), «Is the top object adequately supported by the bottom object? Young infants' understanding of support relations», Cognitive Development, 5, 29-54.

BAILLARGEON, R.; SPELKE, E. S., y WASSERMAN, S. (1985), «Object permanence in five-month-old infants», Cognition, 20, 191-208.

BAINBRIDGE, W. S. (1997), The Sociology of Religious Movements. Nueva York: Routledge.

BARA, B.; BUCCIARELLI, M., y JOHNSON-LAIRD, P. N. (1995), «The development of syllogistic reasoning», American Journal of Psychology, 108, 157-193.

BARGH, J. A.; CHEN, M., y BURROWS, L. (1996), «Automaticity of social behavior: direct effects of trait construct and stereotype activation on action», Journal of Personality and Social Psychology, 71, 230-244.

BAR-HILLEL, M., y FALK, R. (1982), «Some teasers concerning conditional probabilities», Cognition, 11, 109-122.

BAR-HILLEL, M., y NETER, E. (1996), «Why are people reluctant to exchange lottery tickets?», Journal of Personality and Social Psychology, 70, 17-27.

BAR-HILLEL, Y. (1967), «Dictionaries and meaning rules», Foundations of Language, 3, 409-414.

BAR-HILLEL, Y. (1969), «Colloquium on the role of formal languages», Foundations of Language, 5, 256-284.

BAR-HILLEL, Y., y CARNAP, R. (1964), «An outline of a theory of semantic information». En Bar- Hillel, Y., Language and Information Processing. Reading, MA: Addison-Wesley.

BARNES, B., y BLOOR, D. (1982), «Relativism, rationalism, and the sociology of knowledge». En Hollis, M., y Lukes, S. (eds.), Rationality and Relativism. Oxford: Basil Blackwell, pp. 21-47.

BARÓN, J. (2000), Thinking and Deciding, 3ª ed. Nueva York: Cambridge University Press.

BARÓN, J., y RITOV, I. (1994), «Reference points and omission bias», Organizational Behavior and Human Decisión Processes, 59, 475-498.

BARON-COHEN, S.; LESLIE, A. M., y FRITH, U. (1985), «Does the autistic child have a “theory of mind”?», Cognition, 21, 37-46.

BARRES, P., y JOHNSON-L.AIRD, P. N. (2003), «On imagining what is true (and what is false)», Thinking&Reasoning, 9, 1-42.

BARROUILLET, P. (1997), «Modifying the representation of if... then sentences in adolescents by inducing a structure mapping strategy», Current Psychology of Cognition, 16, 609-637.

BARROUILLET, P., y LECAS, J-F. (1998), «How can mental models theory account for content effects in conditional reasoning? A developmental perspective», Cognition, 209-253.

BARROUILLET, P., y LECAS, J.-F. (1999), «Mental models in conditional reasoning and working memory», Thinking & Reasoning, 5, 289-302.

BARROUILLET, P.; GROSSET, N., y LECAS, J. F. (2000), «Conditional reasoning by mental models: chronometric and developmental evidence», Cognition, 75, 237-266.

BARSALOU, L. W. (1983), «Ad hoc categories», Memory & Cognition, 11, 211-227.

BARSALOU, L. W. (1999), «Perceptual symbol systems», Behavioral and Brain Sciences, 22, 577-660.

BARWISE, J. (1986), «Conditionals and conditional information». En Traugott, E. C.; ter Meulen, A.; Reilly, J. S., y Ferguson, C. A. (eds.), On Conditionals. Cambridge: Cambridge University Press, pp. 21-54.

BARWISE, J. (1987), The Situation in Logic. Stanford, CA: CSLI Publications.

BARWISE, J. (1993), «Everyday reasoning and logical inference», Behavioral and Brain Sciences, 16, 337-338.

BARWISE, J., y COOPER, R. (1981), «Generalized quantifiers and natural language», Linguistics and Philosophy, 4, 159-219.

BARWISE, J., y ETCHEMENDY, J. (1987), The Liar: an Essay in Truth and Circularity. Nueva York: Oxford University Press.

BARWISE, J., y ETCHEMENDY, J. (1992), «Hyperproof: logical reasoning with diagrams». En Narayanan, N. H. (ed.), AAA1 Spring Symposium on Reasoning with Diagratnmatic Representations. March 25-27th, Stanford University, California, pp. 80-84.

BARWISE, J., y ETCHEMENDY, J. (1994), Hyperproof. Stanford, CA: Center for the Study of Language and Information.

BATESON, G. (1979), On Mind and Nature. Londres: Wildwood House.

BAUER, M. I., y JOHNSON-LAIRD, P. N. (1993), «How diagrams can improve reasoning». Psychological Science, 4, 372-378.

BEARDSLEY, M. C. (1950), Practical Logic. Englewood Cliffs, NJ: Prentice-Hall.

BECHARA, A.; DAMASIO, A. R.; DAMASIO, H., y ANDERSON, S. W. (1994), «Insensitivity to future consequences following damage to human prefrontal cortex», Cognition, 50, 7-15.

BECK, A. T. (1976), Cognitive Therapy and the Emotional Disorders. Nueva York: Meridian.

BECK, A. T. (1991), «Cognitive therapy: a 30-year retrospective», American Psychologist, 46, 368-375.

BECK, A. T.; RUSH, A.; SHAW, B., y EMERY, G. (1979), Cognitive Therapy of Depression. Nueva York: Guilford Press.

BEDFORD, F. L. (1997), «False categories in cognition: the not-the-liver fallacy», Cognition, 64, 231-248.

BEGG, I., y DENNY, J. (1969), «Empirical reconciliation of atmosphere and conversión interpretations of syllogistic reasoning», Journal of Experimental Psychology, 81, 351-354.

BELL, V. (1999), «The model method: an aid to improve reasoning». Presentación de una tesis doctoral sin publicar, Departamento de Psicología. Universidad de Princeton.

BELL, V., y JOHNSON-LAIRD, P. N. (1998), «A model therory of modal reasoning», Cognitive Science, 22, 25-51.

BENNETT, J. F. (2003), A Philosophal Guide to Conditionals. Nueva York: Oxford University Press.

BENTHAM, J. (1996), Burns, J. H., y Hart, H. L. A. (eds.), An intrtoduction to the Principles of Morals and Legislation. Oxford: Clarendon Press. (Publicado originalmente en 1789.)

BERARDI-COLETTA, B.; BUYER, L.; DOMINOWSKI, R., y RELLINGER, E. (1995), «Metacognition and problem-solving: a process-oriented approach», Journal of Experimental Psychology: Learning, Memory and Cognition, 21, 205-223.

BERMOND, B.; NIEUWENHUYSE, B.; FASOTTI, L., y SCHUERMAN, J. (1991), «Spinal cord lesions, peripheral feedback, and intensities of emotional feelings», Cognition and Emotion, 5, 201-220.

BERRY, G. D. W. (1952), «Peirce's contributions to the logic of statements and quantifiers». En Wiener, P., y Young, F. (eds.), Studies in the Philosophy of Charles S. Peirce. Cambridge, MA: Harvard University Press.

BERWICK, R. C. (1986), «Learning from positive-only examples: the Subset principie and three case studies». En Michalski, R. S.; Carbonell, J. G., y Mitchell, T. M. (eds.), Machine Learning: an Artificial Inteligence Approach, Vol. II. Los Altos, CA: Morgan Kaufmann.

BETH, E. W. (1971), Aspects of Modern Logic. Dordecht: Reidel.

BETH, E. W., y PIAGET, J. (1966), Mathematical Epistemology and Psychology. Dordrecht: Reidel.

BIRBAUMER, N.; GRODD, W.; DIEDRICK, O.; CLOSE, U.; ERB, M.; LOTZE, M.; SCHNEIDER, F.; WEISS, U., y FLOR, H. (1998), «FMRI reveals amygdala activation to human faces in social phobics», NeuroReport, 9, 1223-1226.

BIRCH, S. A. J. (2005), «When knowledge is a curse: children's and adults' reasoning about mental states», Current Directions in Psychological Science, 14, 25-29.

BIRNEY, D., y HALFORD, G. S. (2002), «Cognitive complexity of suppositional reasoning: an application of relational complexity to the knight-knave task», Thinking and Reasoning, 8, 109-134.

BLACK, A.; FREEMAN, P., y JOHNSON-LAIRD, P. N. (1986), «Plausibility and the comprehension of text», British Journal of Psychology, 77, 51-62.

BLACKBURN, S. (2005), Truth: a Guide. Oxford: Oxford University Press.

BLANCHETTE, I., y RICHARDS, A. (2004), «Reasoning about emotional and neutral materials: is logic affected by emotion?», Psychological Science, 15, 745-752.

BLANTON, H., y JACCARD, J. (2006), «Arbitrary metrics in psychology», American Psychologist, 61, 27-41.

BLOOM, P., y GERMAN, T. P. (2000), «Two reasons to abandon the false belief task as a test of theory of mind», Cognition, 77, B25-B31.

BODEN, M. A. (1980), Jean Piaget. Nueva York: Viking.

BOLTZMANN, L. (1974), «On the fundamental principies and equations of mechanics». En McGuiness, B. (ed.), Ludwig Boltzmann: Theoretical Physics and Philosophical Prohlems. Boston: Reidel, pp. 101-128. (Originalmmente publicado en 1899.)

BONNEFON, J.-F. (2004), «Reinstatement, floating conclusions, and the credulity of mental model reasoning», Cognitive Science, 28, 621-631.

BOOLE, G. (1854), An Investigation of the Laws of Thought on Which are Founded the Mathematical Theories of Logic and Probahilities. Londres: Macmillan. [Ed. cast.: Investigación sobre las leyes del pensamiento, trad. J. A. Suárez Hernández, Paraninfo (1982).]

BOOLE, M. E. (1909), Philosophy and Fun of Algebra. Londres: C. W. Daniel.

BOOLOS, G., y JEFFREY, R. (1989), Computability and Logic, 3ª ed. Cambridge: Cambridge University Press.

BORGES, J. L. (1997), Ficciones. Alianza Editorial.

BOUDREAU, G., y PIGEAU, R. (2001), «The mental representation and processes of spatial deductive reasoning with diagrams and sentences», International Journal of Psychology, 36, 42-52.

BOUQUET, P., y WARGLIEN, M. (1999), «Mental models and local model semantics: the problem of information integration». En Bagnara, S. (ed.), European Conference on Cognitive Science, 1999. Siena: Universitá degli Studi di Siena, pp. 169-178.

BOWER, G. H.; BLACK, J. B., y TURNER, T. J. (1979), «Scripts in memory for text», Cognitive Psychology, 11, 177-220.

BOWKER, G. (2003), George Orwell. Londres: Little, Brown.

BOWLBY, J. (1988), A Secure Base: Clinical Applications of Attachment Theory. Londres: Routledge.

BOWMAN, W. E. (1956), The Ascent of Rum Doodle. Londres: Max Parrish.

BOZZI, P. (1991), Física Ingenua. Milán, Italia: Garzanti.

BRADSHAW, G. (1992), «The airplane and the logic of invention». En Giere, R. N. (ed.), Cognitive Models of Science, Vol. XV. Minneapolis: Universidad of Minnesota Press, pp. 239-250.

BRAINE, M. D. S. (1978), «On the relation between the natural logic of reasoning and standard logic», Psychological Review, 85, 1-21.

BRAINE, M. D. S., y O'BRIEN, D. P. (1991), «A theory of If: a lexical entry, reasoning program, and pragmatic principles», Psychological Review, 98, 182-203.

BRAINE, M. D. S., y O'BRIEN, D. P. (eds.) (1998), Mental Logic. Mahwah, NJ: Lawrence Erlbaum Associates.

BRAINE, M. D. S., y RUMAIN, B. (1983), «Logical reasoning». En Flavell, J. H., y Markman, E. M. (eds.), Carmichael's Handbook of Child Psychology, Vol. III. Cognitive Development, 4ª ed. Nueva York: Wiley, pp. 263-340.

BRAINE, M. D. S., REISER, B. J., y RUMAIN, B. (1984), «Some empirical justification for a theory of natural propositional logic», The Psychology of Learning and Motivation, Vol. 18. San Diego, CA: Academic Press, pp. 313-371.

BREITER, H. C.; RAUCH, S. L.; KWONG, K. K.; BAKER, J. R.;

WEISSKOFF, R. M.; KENNEDY, D. N.; KENDRICH, A. D.; DAVIS, T. L.; JIANG, A.; COHEN, M. S., STERN, C. E.; BELLIVEAU, J. W.; BAER, L.; O'SULLIVAN, R. L.; SAVAGE, C. R.; JENIKE, M. A., y ROSEN, B. R. (1996), «Functional magnetic resonance imaging in symptom provocation in obsessive-compulsive disorder», Archives of General Psychiatry, 53, 595-606.

BREWKA, G.; DIX, J., y KONOLIGE, K. (1997), Nonmonotonic Reasoning: an Overview. Stanford, CA: CLSI Publications, Universidad de Stanford.

BROADBENT, D. E., y BROADBENT, M. (1988), «Anxiety and attentional bias: state and trait», Cognition and Emotion, 2, 165-183.

BROADBENT, D. E.; FITZGERALD, P., y BROADBENT, M. H. P. (1986), «Implicit and explicit knowledge in the control of complex systems», British Journal of Psychology, 17, 33-50.

BRUNER, J. S.; GOODNOW, J. A., y AUSTIN, G.A. (1956), A Study of Thinking. Nueva York: Wiley.

BRYANT, D. J. (1992a), «A spatial representation system in humans», Psycoloquy, 3(16), Space.l [Revista on-line].

BRYANT, D. J. (1992b), «More on the spatial representation system», Psycoloquy, 3, (44), Space.5 [Revista on-line].

BUCCIARELLI, M., y JOHNSON-LAIRD, P. N. (1999), «Strategies in syllogistic reasoning», Cognitive Science, 23, 247-303.

BUCCIARELLI, M., y JOHNSON-I.AIRD, P. N. (2000), «Is there an innate module for deontic reasoning?». En García-Madruga, J.; Carriedo, N., y González-Labra, M. J. (eds.), Mental Models in Reasoning. Madrid: UNED, pp. 227-239.

BUCCIARELLI, M., y JOHNSON-LAIRD, P. N. (2001), «Falsification and the role of the theory of mind in the reduced array selection task», Current Psychology Letters: Behavior, Brain and Cognition, 4, 7-22.

BUCCIARELLI, M., y JOHNSON-LAIRD, P. N. (2005), «Naïve deontics: a theory of meaning, representation, and reasoning», Cognitive Psychology, 50, 159-193.

BUCHANAN, T. W., y ADOPHS, R. (2002), «The role of the human amygdala in emotional modulation of long-term declarative memory». En Moore, S. C., y Oaksford, M. (eds.), Ernotional Cognition: from Brain to Behaviour. Amsterdam: John Benjamins, pp. 9-34.

BUDIANSKY, S. (2000), Battle of Wits: the Complete Story of Codebreaking in World War II. Nueva York: Simón & Schuster.

BULBROOK, M. E. (1932), «An experimental study into the existence and nature of "insight"», American Journal of Psychology, 44, 409-453.

BULLER, D. J. (2005), Adapting Minds: Evolution and the Persistent Quest for Human Nature. Cambridge, MA: MIT Press.

BULLOCK, M.; GELMAN, R., y BAILLARGEON, R. (1982), «The development of causal reasoning». En Friedman, W. J. (ed.), The Developmental Psychology of Time. Orlando, FL: Academic Press, pp. 209-254.

BURKE, P. (1986), «Strengths and weaknesses in the history of mentalities», History of European Ideas, 7, 439-451.

BYRNE, R. M. J. (1989), «Suppressing valid inferences with conditionals». Cognition, 31, 61-83.

BYRNE, R. M. J. (1991), «Can valid inferences be suppressed?», Cognition, 39, 71-78.

BYRNE, R. M. J. (1996), «A model theory of imaginary thinking». En Oakhill, J., y Garnham, A. (eds.), Mental Models in Cognitive Science. Hove: Erlbaum (UK), Taylor & Francis, pp. 155-174.

BYRNE, R. M. J. (1997), «Cognitive processes in counterfactual thinking about what might have been». En Medin, D. L. (ed.), The Psychology of Learning and Motivation, Vol. 37. San Diego, CA: Academic Press, pp. 105-154.

BYRNE, R. M. J. (2002), «Mental models and counterfactual thoughts about what might have been», Trends in Cognitive Sciences, 6, 426-431.

BYRNE, R. M. J. (2005), The Rational Imagination: How People Create Alternatives to Reality. Cambridge, MA: MIT Press.

BYRNE, R. M. J., y HANDLEY, S. J. (1992), «Reasoning strategies», Irish Journal of Psychology, 13, 111-124.

BYRNE, R. M. J., y HANDLEY, S. J. (1997), «Reasoning strategies for suppositional deductions», Cognition, 62, 1-49.

BYRNE, R. M. J., y JOHNSON-LAIRD, P. N. (1989), «Spatial reasoning», Journal of Memory and Language, 28, 564-575.

BYRNE, R. M. J., y MCELENEY, A. (2000), «Counterfactual thinking about actions and failures to act», Journal of Experimental Psychology: Learning, Memory, and Cognition, 26, 1318-1331.

BYRNE, R. M. J., y TASSO, A. (1999), «Deductive reasoning with tactual, possible, and counterfactual conditionals», Memory & Cognition, 27, 726-740.

BYRNE, R. M. J.; HANDLEY, S. J., y JOHNSON-LAIRD, P. N. (1995), «Reasoning from suppositions», Quarterly Journal of Experimental Psychology, 48A, 915-944.

BYRNE, R. M. J.; ESPINO, O., y SANTAMARÍA, C. (1999), «Counterexamples and the suppression of inferences», Journal of Memory and Language, 40, 347-373.

BYRNES, J. P., y DUFF, M. A. (1989), «Young children's comprehension of modal expressions», Cognitive Development, 4, 369-387.

CALDER, A. J.; YOUNG, A. W.; ROWLAND, D.; PERRETT, D.; HODGES, J. R., y ETCOFF, N. L. (1996), «Facial emotion recognition after bilaterial amygdala damage: difterentially severe impairment of fear», Cognitive Neuropsychology, 13, 699-745.

CAMPBELL, D. (1960), «Blind variation and selective retention in creative thought as in other knowledge processes», Psychological Review, 67, 380-400.

CARDACI, M.; GANGEMI, A.; PENDOLINO, G., y DI NUOVO, S. (1996), «Mental models vs. integrated models: explanations of syllogistic reasoning», Perceptual and Motor Skills, 82, 1377-1378.

CAREY, S. (1985a), «Are children fundamentally different kinds of thinkers and

learners than adults?». En Chipman, S. F.; Segal, J. W., y Glaser, R. (eds.), Thinking and Learning Skills, Vol. 2: Research and Open Questions. Hillsdale, NJ: Erlbaum. pp. 485-517.

CAREY, S. (1985b), Conceptual Change in Childhood. Cambridge: MIT Press.

CAREY, S. (1988), «Conceptual differences between children and adults», Mind & Language, 3, 167-181.

CAREY, S. (1991), «Knowledge acquisition: enrichment or conceptual change?». En Carey, S., y Gelman, R. (eds.), The Epigenesis of Mind: Essays on Biology and Cognition. Hillsdale, NJ: Erlbaum, pp. 257-291.

CARLISLE, E., y SHAFIR, E. (2005), «Questioning the cheater-detection hypothesis: new studies with the selection task», Thinking & Reasoning, 11, 97-122.

CARNAP, R. (1966), Philosophical Foundations of Physics. Nueva York: Basic Books.

CARREIRAS, M., y SANTAMARÍA, C. (1997), «Reasoning about relations: spatial and nonspatial problems», Thinking & Reasoning, 3, 191-208.

CARROLL, L. (1958), Symbol Logic and the Game of Logic. (Dos libros reunidos en uno.) Nueva York: Dover.

CASE, R. (1985a), «A developmentally based approach to the problem of instructional design». En Chipman, S. F.; Segal, J. W., y Glaser, R. (eds.), Thinking and Learning Skills, Vol. 2: Research and Open Questions. Hillsdale, NJ: Erlbaum, pp. 545-562.

CASE, R. (1985b), Intellectual Development: Birth to Adulthood. Nueva York: Academic Press.

CHAITIN, G. J. (1998), The Limits of Mathematics: a Course on Information Theory and the Limits of Formal Reasoning. Singapuo: Springer-Verlag.

CHALMERS, D. J.; FRENCH, R. M., y HOFSTADTER, D. R. (1992), «High-level perception, representation, and analogy: a critique of artificial intelligence methodology», Journal of Experimental & Theoretical Artificial Intelligence, 4,

185-211.

CHAPMAN, L. J., y CHAPMAN, A. P. (1959), «Atmosphere effect re-examined», Journal of Experimental Psychology, 58, 220-226.

CHATER, N., y OAKSFORD, M. (1999), «The probability heuristics model of syllogistic reasoning», Cognitive Psychology, 38, 191-258.

CHATER, N.; HEIT, E., y OAKSFORD, M. (2005), «Reasoning». En Lamberts, K., y Goldstone, R. L. (eds.), Handbook of Cognition. Thousand Oaks, CA: Sage, pp. 297-320.

CHENG, P. W. (1997), «From covariation to causation: a causal power theory», Psychologcal Review, 104, 367-405.

CHENG, P. W., y HOLYOAK, K. J. (1985), «Pragmatic reasoning schemas», Cognitive Psychology, 17, 391-416.

CHENG, P. W., y HOLYOAK, K. J. (1989), «On the natural selection of reasoning theories», Cognition, 33, 285-313.

CHENG, P. W., y NOVICK, L. R. (1990), «A probabilistic constrast model of causal induction», Journal of Personality and Social Psychology, 58, 545-567.

CHENG, P. W., y NOVICK, L. R. (1991), «Causes versus enabling conditions», Cognition, 40, 83-120.

CHENG, P. W.; HOLYOAK, K. J.; NISBETT, R. E., y OLIVER, L. M. (1986), «Pragmatic versus syntactic approaches to training deductive reasoning», Cognitive Psychology. 18, 293-328.

CHERNIAK, C. (1986), Minimal Rationality. Cambridge, MA: MIT Press.

CHERRY, C. (1980), On Human Comtnunication: a Revíew, a Survey, and a Criticism, 3ª ed. (edición revisada). Cambridge, MA: MIT Press.

CHERUBINI, P., y JOHNSON-LAIRD, P. N. (2004), «Does everyone love everyone? The psychologyof iterative reasoning», Thinking & Reasoning, 10, 31-53.

CHERUBINI, P.; GARNHAM, A.; OAKHILL, J., y MORLEY, E. (1998), «Can any ostrich fly? Some new data on belief bias in syllogistic reasoning», Cognition, 69, 179-218.

CHI, M. T. H. (1976), «Short-term memory limitations in children: capacity or processing deficits?», Memory & Cognition, 4, 559-572.

CHI, M. T. H. (1978), «Knowledge structures and memory development». En Siegler, R. S. (ed.), Children's Thinking: What Develops? Hillsdale, NJ: Erlbaum, pp. 73-96.

CHOMSKY, N. (1995), The Minimalist Program. Cambridge, MA: MIT Press. [Ed. cast.: El programa minimalista, trad. de J. Romero Morales, Madrid, Alianza Editorial, 1999.]

CHRISTIE, A. (1960), Murder on the Orient Express. Nueva York: Penguin Putnam. (Publicada originalmente en 1933.)

CHRONICLE, E. P.; ORMEROD, T. C., y MACGREGOR, J. N. (2001), «When insight just won't come: the failure of visual cues in the nine-dot problem», Quarterly Journal of Experimental Psychology, 54A, 903-919.

CLARK, H. H. (1969), «Linguistic processes in deductive reasoning», Psychological Review, 76, 387-404.

CLARK, H. H. (1973), «Space, time, semantics, and the child». En Moore, T. E. (ed.), Cognitive Development and the Acquisition of Language. Nueva York: Academic Press, pp. 28-64.

CLARK, H. H., y CHASE, W. G. (1972), «On the process of comparing sentences against pictures», Cognitive Psychology, 3, 472-517.

CLARK, H. H., y HAVILAND, S. E. (1977), «Comprehension and the given-new contract». En Freedle, R. O. (ed.), Discourse Processes: Advances in Research and Theory, Vol. 1: Discourse Production and Comprehension. Norwood, NJ: Ablex, pp. 1-40.

CLEMENT, C. A., y FALMAGNE, R. J. (1986), «Logical reasoning, world knowledge, and mental imagery: interconnections in cognitive processes», Memory & Cognition, 14, 299-307.

COHEN, E. D. (2003), Self-control through the Power of Reason: What would Aristotle do? Nueva York: Prometheus Books.

COHEN, L. J. (1981), «Can human irrationality be experimentally demonstrated?», Behavioral and Brain Sciences, 4, 317-370.

COHEN, M. R., y NAGEL, E. (1934), An Introduction to Logic and Scientific Method. Londres: Routledge & Kegan Paul.

COLE, M.; GAY, J.; GLICK, J. A., y SHARP, D. W. (1971), The Cultural Context of Learning and Thinking. Nueva York: Basic Books.

COLLINS, A., y MICHALSKI, R. (1989), «The logic of plausible reasoning: a core theory», Cognitive Science, 13, 1-49.

COMBS, H. (1979), Kill Devil Hill: Discovering the Secret of the Wright Brothers. (Con M. Caidin.) Boston, MA: Houghton Mifflin.

CONAN DOYLE, A. (1930), The Complete Sherlock Holmes, Vols. I y II. Nueva York: Doubleday.

CONSUMER REPORTS (1995), Mental Health: Does Therapy Help? Noviembre, 734-739.

COOK, V. J., y NEWSON, M. (1996), Chomsky's Universal Grammar: an Introduction, 2ª ed. Oxford: Blackwell.

COPELAND, D. E., y RADVANSKY, G. A. (2004), «Working memory and syllogistic reasoning», Quarterly Journal of Experimental Psychology, 57A, 1437-1457.

CORLETT, P. R.; AITKEN, M. R. F.; DICKINSON, A.; SHANKS, D. R.; HONEY, G. D.; HONEY, R. A. E.; ROBBINS, T. W.; BULLMORE, E. T., y FLETCHER, P. C. (2004), «Prediction error during retrospective revaluation of causal associations in humans: fMRl evidence in favor of an associative model of learning», Neuron, 44, 877-888.

COSMIDES, L. (1989), «The logic of social exchange: Has natural selection shaped how humans reason?», Cognition, 31, 187-276.

COSMIDES, L., y TOOBY, J. (1996), «Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty», Cognition, 58, 1-73.

COZZUTTI, C., y VALLORTIGARA, G. (2001), «Hemispheric memories for the content and position of food caches in the domestic chick», Behavioral Neuroscience, 115, 305-313.

CRAIG, D. L.; NERSESSIAN, N. J., y CATRAMBONE, R. (2002), «Perceptual simulation in analogical problem solving». En Magnani, L., y Nersessian, N. (eds.), Model-Based Reasoning: Science, Technology, Values. Nueva York: Kluwer Academic/Plenum, pp. 169-187.

CRAIK, K. (1943), The Nature of Explanation. Cambridge: Cambridge University Press.

CRICK, F. (1988), What Mad Pursuit. Nueva York: Basic Books.

CROUCH, T. D. (1989), The Bishop's Boys: a Life of Wilhur and Orville Wright. Nueva York: W.W. Norton.

CRUTCHFIELD, R. S. (1955), «Conformity and character», American Psychologist, 10, 191-198.

CUMMINS, D. D. (1995), «Naïve theories and causal deduction», Memory & Cognition, 23, 646-658.

CUMMINS, D. D. (1998), «The pragmatics of causal inference», Proceedings of the Twentieth Annual Conference of the Cognitive Science Society, p. 9.

CUMMINS, D. D.; LUBART, T.; ALKSNIS, O., y RIST, R. (1991), «Conditional reasoning and causation», Memory & Cognition, 19, 274-282.

CUNNINGHAM, W. A.; JOHNSON, M. K.; RAYE, C. L.; GATENBY, J. C.; GORE, J. C., y BANAJI, M. R. (2004), «Separable neural components in the processing of Black and White faces», Psychological Science, 15, 806-813.

CURTISS, S. (1977), Genie: a Psycholinguistic Study of a Modern Day "Wild Child". Nueva York: Academic Press.

DADLEZ, E. M. (1997), What's Hecuba to Him?Fictional Events and Actual Emotions. University Park, PA: Pennsylvania State University Press.

DAMASIO, A. R. (1994), Descartes' Error: Emotion, Reason, and the Human Brain. Nueva York: Grosset/Putnam. [Ed. cast.: El error de Descartes, trad. de J. Ros, Barcelona, Crítica, 1996.]

DARWIN, C. (1887), The Life and Letters of Charles Darwin. Darwin, F. (ed.), Londres: Murray.

DARWIN, C. (1872), The Expression of the Emotions in Man and Animals. Nueva York: Philosophical Library. [Ed. cast.: La expresión de las emociones en los animales y en el hombre, trad. de R. Fernández Rodríguez, Madrid, Alianza Editorial, 1998.]

DARWIN, C. (1839), Voyage of the Beagle. Nueva York: Penguin Putman. [Ed. cast.: El viaje del Beagle, trad. de O. Albó, Pearson Alhambra, 1985.]

DASEN, P. (1994), «Culture and cognitive development from a Piagetian perspective». En Lonner, W., y Malpass, R. (eds.), Psychology and Culture. Boston: Allyn y Bacon, pp. 145-149.

DAVIDSON, D. (1980), «Actions, reasons and causes». En Davidson, D. (ed.), Essays on Actions and Events. Oxford: Oxford University Press, pp. 3-20.

DAVIDSON, R. J.; PIZZAGALLI, D.; NITSCHKE, J. B., y PUTMAN, K. (2002), «Depression: Perspective from affective neuroscience», Annual Review of Psychology, 53, 545-574.

DAWES, R. (2001), Everyday Irrationality: How Pseudoscientists, Lunatics, and the Rest of Us Fail to Think Rationally. Boulder, CO: Westview Press.

DAWKINS, R. (1976), The Selfish Gene. Oxford: Oxford University Press.

DAY, C. (1996), «Understanding of the French modal verbs "pouvoir" and "devoir" in school children and adults», Current Psychology of Cognition, 15, 535-553.

DEARY, I. J. (2005), «Intelligence, health and death», The Psychologist, 18, 610-613.

DE BONO, E. (1992), Serious Creativity: Using the Power of Lateral Thinking to Create New Ideas. Londres: Harper Collins. [Ed. cast.: El pensamiento creativo: el poder del pensamiento lateral para la creación de nuevas ideas, trad. de Ofelia Castillo, Barcelona, Paidós Ibérica, 1995.]

DEHAENE, S.; IZARD, V.; PICA, P., y SPELKE, E. (2006), «Core knowledge of geometry in an Amazonian indigene group», Science, 311, 381-384.

DELIUS, J. D., y SIEMANN., M. (1998), «Transitive responding in animals and humans. Exaptation rather trian adaptation?», Behavioural Processes, 42, 107-137.

DELVAL, J. A., y RIVIERE, A. (1975), «"Si llueve, Elisa lleva sombrero": Una investigación psicología sobre la tabla de verdad del condicional», Revista De Psicología General y Aplicada, 136, 825-850.

DESCARTES, R. (1637), El discurso del método. En Haldane, E. S., y Ross, G. R. T. (trans.), The Philosophical Works of Descartes, 2 Vols. Cambridge: Cambridge University Press.

DESCARTES, R. (1984-91), «Les passions de l'âme (The Passions of the Soul)». En Cottingham, J.; Stoothoff, R.; Murdoch, D., y Kenny, A. (eds. y trad.), The Philosophical Writings of Descartes, Vol. 1. Cambridge: Cambridge University Press. (Publicado originalmente en 1649.)

DE SOTO, C.; LONDRES, M., y HANDEL, S. (1965), «Social reasoning and spatial paralogic», Journal of Personality and Social Psychology, 2, 513-521.

DEWDNEY, A. K. (1989), The Turing Ómnibus: 61 Excursions in Computer Science. Rockville, MD: Computer Science Press.

DIAS, M.; ROAZZI, A., y HARRIS, P. L. (2005), «Reasoning from unfamiliar premises: a study with unschooled adults», Psychological Science, 16, 550-554.

DICHEVA, D., y CLOSE, J. (1996), «Mental models of recursion», Journal of Educational Computing Research, 14, 1-23.

DICKSTEIN, L. S. (1975), «Effects of instructions and premise order on errors in syllogistic reasoning», Journal of Experimental Psychology: Human Learning and Memory, 104, 376-384.

DIEUSSAERT, K.; SCHAEKEN, W.; DE NEYS, W., y D'YDEWALLE, G. (2000), «Initial belief state as a predictor of belief revisión», Current Psychology of Cognition, 19, 277-288.

DIJKSTERHUIS, A.; BOS, M. W.; NORDGREN, L. F., y VAN BAAREN, R. B. (2006), «On making the right choice: the deliberation-without-attention effect», Science, 311, 1005-1007.

DIXIT, A. K., y NALEBUFF, B. J. (1991), Thinking Strategically: the Competitive Edge in Business Politics, and Everyday Life. Nueva York: Norton.

DOMINOWSKI, R. L. (1995), «Content effects in Wason's selection task». En Newstead, S. E., y Evans, J. St. B. T. (eds.), Perspectives on Thinking and Reasoning: Essays in Honor of Petcr Wason. Hillsdale, NJ: Erlbaum, pp. 41-65.

DÖRNER, D. (1996), The Logic of Failure: Why Things Go Wrong and What We Can Do to Make Them Right. Nueva York: Holt. (Publicado originalmente en 1989.)

DOSTOYEVSKY, F. M. (1864), Notes form Underground. Coulson, J. (trad.), Harmondsworth, Middlesex: Penguin. (Publicado originalmente en 1864.)

DOYLE, J. (1979), «A truth maintenance system», Artificial Intelligence, 12, 231-272.

DREYFUS, H. L. (1972), What Computers Can't Do: a Critique of Artificial Intelligence. Nueva York: Harper y Row.

DSM-IV-TR (2000), American Psychiatric Assocation.

DUDMAN, V. H. (1988), «Indicative and subjunctive conditionals», Analysis, 48, 113-122.

DUNBAR, K. (1995), «How scientists really reason: scientific reasoning in real-world laboratories». En Sternberg, R. J., y Davidson, J. (eds.), The Nature of lnsight. Cambridge, MA: MIT Press, pp. 365-395.

DUNCKER, K. (1945), «On problem solving», Psychological Monographs, 58(5), número completo 270.

DUNNE, P. E. (1988), The Complexity of Boolean Networks. Londres: Academic Press.

DYSON, G. B. (1997), Darwin among the Machines: the Evolution of Global Intelligence. Reading, MA: Helix Books, Addison-Wesley.

EDWARDS, A. W. F. (2004), Cogwheels of theMind: the Story of Venn Diagrams. Baltimore: Johns Hopkins University Press.

EGAN, D. E., y GRIMES-FARROW, D. D. (1982), «Differences in mental representations spontaneously adopted for reasoning», Memory & Cognition, 10, 297-307.

EELLS, E., y SKYRMS, B. (eds.), (1994), Probability and Conditionals. Cambridge: Cambridge University Press.

EICHENBAUM, H. (1999), «The hippocampus». En Wilson, R. A., y Keil, F. C. (eds.), The MIT Encydopedia of the Cognitive Sciences. Cambridge, MA: MIT Press, pp. 377-378.

EILAM, E. (2005), Reversing: Secrets of Reverse Engineering. Indianapolis, IN: Wiley.

EINHORN, H. J., y HOGARTH, R. M. (1986), «Judging probable cause», Psychological Bulletin, 99, 3-19.

EKMAN, P. (1980), «Biological and cultural contributions to body and facial movement in the expression of emotion». En Rorty, A. O. (ed.), Explaining Emotion. Berkeley: University of California Press.

EKMAN, P. (1992a), «An argument for basic emotions», Cognition and Emotion, 6, 169-200.

EKMAN, P. (1992b), «Facial expression of emotion: new findings, new questions», Psychological Science, 3, 34-38.

EKMAN, P. (1995), Telling Lies: Clues to Deceit in the Marketplace, Politics, and Marriage. Nueva York: Norton.

EKMAN, P.; DAVIDSON, R. J.; RICARD, M., y WALLACE, B. A. (2005),

«Buddhist and psychological perspectives on emotions and well-being», Current Directions in Psychological Science, 14, 59-63.

ELEY, T. C., y PLOMIN, R. (1997), «Genetic analyses of emotionality», Current Opinion in Neurobiology, 7, 279-284.

ELIO, R., y PELLETIER, F. J. (1997), «Belief change as propositional update», Cognitive Science, 21, 419-460.

ELLIS, A. (1962), Reason and Emotion in Psychotherapy. Nueva York: Lyle Stuart.

ELSTER, J. (1999), Alchemies ofthe Mind. Cambridge: Cambridge University Press.

ERDELYI, M. H. (1985), Psychoanalysis: Freud's Cognitive Psychology. Nueva York: Freeman.

ERICKSON, J. R. (1974), «A set analysis theory of behavior in formal syllogistic reasoning tasks». En Solso, R. (ed.), Loyola Symposium on Cognition, Vol. 2. Hillsdale, NJ: Erlbaum, pp. 305-329.

ERICSSON, K. A., y SIMON, H. A. (1980), «Verbal reports as data», Psychiological Review, 87, 215-251.

ESPINO, O.; SANTAMARÍA, C., y GARCÍA-MADRUGA, J. A. (2000a), «Activation of end terms insyllogistic reasoning», Thinking & Reasoning, 6, 67-89.

ESPINO, O.; SANTAMARÍA, C.; MESEGUER, E., y CARREIRAS, M. (2000b), «Eye movements during syllogistic reasoning». En García-Madruga, J. A.; Carriedo, N., y González-Labra, M. J. (eds.), Mental Models in Reasoning. Madrid: Universidad Nacional de Educación a Distancia, pp. 179-188.

ESPINO, O.; SANTAMARÍA, C.; MESEGUER, E., y CARREIRAS, M. (2005), «Early and late processes in syllogistic reasoning: evidence from eye-movements», Cognition, 98, B1-B9.

EVANS, G. (1980), «Pronouns», Linguistic Inquiry, 11, 337-362.

EVANS, J. St. B. T. (1972a), «Deductive reasoning and linguistic usage (with special reference to negation)». Tesis doctoral no publicada, Universidad de Londres.

EVANS, J. St. B. T. (1972b), «Interpretation and matching bias in a reasoning task», Quarterly Journal of Experimental Psychology, 24, 193-199.

EVANS, J. St. B. T. (1984), «Heuristic and analytic processes in reasoning», British Journal of Psychology, 75, 451-468.

EVANS, J. St. B. T. (1989), Bias in Human Reasoning; Causes and Consequences. Hillsdale, NJ: Erlbaum.

EVANS, J. St. B. T. (2003), «In two minds: dual process accounts of reasoning», Trends in Cognitive Sciences, 7, 454-459.

EVANS, J. St. B. T., y OVER, D. E. (1996), Rationality and Reasoning. Hove, East Sussex: Psychology Press.

EVANS, J. St. B. T., y OVER, D. E. (2004), If. Oxford: Oxford University Press.

EVANS, J. St. B. T.; BARSTON, J. L., y POLLARD, P. (1983), «On the conflict between logic and belief in syllogistic reasoning», Memory and Cognition, 11, 295-306.

EVANS, J. St. B. T.; NEWSTEAD, S. E., y BYRNE, R. M. J. (1993), Human Reasoning: the Psychology of Deduction. Hillsdale, NJ: Erlbaum.

EVANS, J. St. B. T.; HANDLEY, S. J.; HARPER, C. N. J., y JOHNSON-LAIRD, P. N. (1999), «Reasoning about necessity and possibility: a test of the mental model theory of deduction», Journal of Experimental Psychology: Learning, Memory, and Cognition, 25, 1495-1513.

EVANS, J. St. B. T.; HANDLEY, S. H.; PERHAM, N.; OVER, D. E., y THOMPSON, V. A. (2000), «Frequency versus probability formats in statistical word problems», Cognition, 77, 197-213.

EVANS, J. St. B. T.; HANDLEY, S. H., y OVER, D. E. (2003), «Conditionals and conditional probability», Journal of Experimental Psychology: Learning, Memory and Cognition, 29, 321-335.

EVANS, J. St. B. T.; OVER, D. E., y HANDLEY, S. J. (2005), «Supposition, extensionality, and conditionals: a critique of the mental model theory of Johnson-Laird and Byrne (2002)», Psychological Review, 112, 1042-1052.

EVANS-PRITCHARD, E. E. (1937), Witchcraft: Grades and Magic Among the Azande. Oxford: Clarendon Press.

EYSENCK, H. J., y EYSENCK, S. B. G. (1975), Manual of the Eysenck Personality Questionnaire. Londres: Hodder & Stoughton.

FALK, R. (1992), «A closer look at the probabilities of the notorious three prisoners», Cognition, 43, 197-223.

FALKENHAINER, B.; FORBUS, K. D., y GENTNER, D. (1989), «The structure mapping engine: algorithm and examples», Artificial Intelligence, 41, 1-63.

FALMAGNE, R. J. (1980), «The development of logical competence». En Kluwe, R. H., y Spada, M. (eds.), Developmental Models of Thinking. Nueva York: Academic Press.

FARAH, M. J.; HAMMOND, K. M.; LEVINE, D. L., y CALVANIO, R. (1988), «Visual and spatial mental imagery: dissociable systems of representation», Cognitive Psychology, 20, 439-462.

FEENEY, A., y HANDLEY, S. J. (2000), «The suppression of q card selections: evidence for deductive inference in Wason's Selection Task», Quarterly Journal of Experimental Psychology, 53ª, 4, 1224-1242.

FELDMAN, J. (2004), «How surprising is a simple pattern? Quantifying "Eureka!"», Cognition, 93, 199-224.

FENKER, D. B.; WALDMANN, M. R., y HOLYOAK, K. J. (2005), «Accessing causal relations in semantic memeory», Memory & Cognition, 33, 1036-1046.

FERGUSON, E. S. (1992), Engineering and the Mind's Eye. Cambridge, MA: MIT Press.

FESTINGER, L.; RIECKEN, H., y SCHACTER, S. (1956), When Prophecy Fails: a Social and Psychological Study of a Modern Group that Predicted the

Destruction of the World. Mineapolis, MN: Universidad de Minnesota.

FEYNMAN, R. P. (1985), Surely You’re joking, Mr. Feynman: Adventures of a Curious Character. (Escrito con Ralph Leighton; Ed. Hutchings, E.) Nueva York: W. W. Norton.

FEYNMAN, R. P. (1988), What Do You Care What Other People Think? Nueva York: Norton.

FEYNMAN, R. P.; LEIGHTON, R. B., y SANDS, M. (1963), The Feynman Lectures on Physics, Vol. 1: Mainly Mechanics, Radiation, y Heat. Reading, MA: Addison-Wesley. [Ed. cast.: Física, trad. de Enrique Oelker, Hugo Espinosa, Carlos A. Heras y Juan Martín, Wilmington, Addison-Wesley Iberoamericana, 1987.]

FIDDICK, L. (2004), «Domains of deontic reasoning: resolving the discrepancy between the cognitive and moral reasonin literatures», Quarterly Journal of Experimental Psychology, 57ª, 447-474.

FIEDLER, K., y SCHENCK, W. (2001), «Spontaneous inferences from pictorially presented behaviors», Personality and Social Psychology Bulletin, 27, 1533-1546.

FILLENBAUM, S. (1976), «Inducements: On the phrasing and logic of conditional promises, threats and warnings», Psychological Research, 38, 231-250.

FILLMORE, C. J. (1971), «Toward a theory of deixis». Artículo presentado en la Coferencia del Pácifico sobre Constractivos Lingüísticos y Universales del Lenguaje, Universidad de Hawaii, Honolulú, enero.

FINDLAY, A. (1937), A Hundred Years of Chemistry. Londres: Duckworth.

FINKE, R. A., y SLAYTON, K. (1988), «Explorations of creative visual synthesis in mental imagery», Memory and Cognition, 16, 252-257.

FISHER, A. (1988), The Logic ofReal Arguments. Cambridge: Cambridge University Press.

FISHER, D. L. (1981), «A three-factor model of syllogistic reasoning: the study

of isolable stages», Memory & Cognition, 9, 496-514.

FISHER, S. C. (1916), «The process of generalizing abstraction; and its product, the general concept», Psychological Monographs, 21, No. 2 (número 90 completo).

FITCH, W. T.; HAUSER, M., y CHOMSKY, N. (2005), «The evolution of the language faculty: clarifications and implications», Cognition, 97, 179-210.

FLAVELL, J. H. (1963), The Developmental Psychology of Jean Piaget. Princeton, NJ: Van Nostrand.

FODOR, J. A. (1975), The Language of Thought. Nueva York: Crowell.

FODOR, J. A. (1980), «Fixation of belief and concept acquisition». En Piattelli-Palmarini, M. (ed.), Lenguage and Learning: the Debate Between Jean Piaget and Noam Chomsky. Londres: Routledge & Kegan Paul, pp. 143-149.

FODOR, J. A. (1981), Representations Brighton, Sussex: Harvester Press.

FODOR, J. A. (1983), The modurlarity of Mind: an Essay on Faculty Psychology. Cambridge, MA: Bradford Books, MIT Press.

FODOR, J. A., y MCLAUGHLIN, B. P. (1990), «Connetcionism and the problem of systematicity: why Smolensky's solution doesn't work», Cognition, 35, 183-204.

FODOR, J. A., y PYLYSHYN, Z. (1988), «Connectionism and cognitive architecture: a critical analysis», Cognition, 35, 3-71.

FOOT, P. (2002), «The problem of abortion ,and the doctrine of the double effect». En Virtues and Vices. Oxford: Oxford University Press, pp. 19-32. (Ensayo publicado originalmente en 1967.)

FORD, M. (1995), «Two modes of mental representation and problem solution in syllogistic reasoning», Cognition, 54, 1-71.

FORKMAN B. (2000), «Domestic hens have declarative representations», Animal Cognition, 3, 135-137.

FORTENBAUGH, W. (1975), Aristotle on Emotions. Londres: Duckworth.

FRAYN, M. (1965), The Tin Men. Boston: Little, Brown.

FRANK, R. (1988), Passions Within Reason: the Strategic Role of Emotions. Nueva York: Norton.

FRANKLIN, J. (2001), The Science of Conjecture: Evidence and Probability Before Pascal. Baltimore, MD: Johns Hopkins University Press.

FRASER, J. G. (1890), The Golden Bough: a Study of Magic and Religion. Oxford: Oxford University Press.

FREEMAN, J. B. (1988), Thinking Logically. Englewood Cliffs, NJ: Prentice-Hall.

FREGE, G. (1967), «Begriffsschrift: a formula language, modeled upon that of arithmetic, for pure thought». En Van Heijenoort, J. (ed.), From Fregel to Gödel: a Source Book in Mathematical Logic, 1879-1931. Cambridge, MA: Harvard University Press, pp. 5-82. (Publicado originalmente en aleman en 1879.)

FREUD, S. (1915), «The Unconscious». En The Standard Edition of the Complete Works of Sigmund Freud, Vol. 14. (Ed. y trad. James Strachey.) Londres: the Hogarth Press and the Institute of Psychoanalys, 1953-74, pp. 161-215.

FREUD, S. (1973), New Introductory Lectures On Psychoanalysis. Harmondsworlh, Middx: Penguin. (Trad. J. Strachey, publicado por primera vez en 1964; versión original alemana publicada en 1933.)

FRIJDA, N. H. (1986), The Emotions. Cambridge: Cambridge University Press.

FROST, W. H. (ed.), (1965), Snow on Cholera, Being a Reprint of Two Papers by John Snow MD Together with a Biographical Memoir by J.W. Richardson. Nueva York: Hafner. Publicado originalmente en 1936.

FUHRMANN, A. (1997), An Essay on Contraction. Stanford, CA: CSLI Publications.

FUJIMURA, K. (1978), The Tokyo Puzzles. Nueva York: Scribner. (Publicado

originalmente en 1884.)

GALOTTI, K. M.; BARON, J., y SABINI, J. P. (1986), «Individual differences in syllogistic reasoning: deduction rules or mental models?», Journal of Experimental Psychology: General, 115, 16-25.

GALTON, F. (1883), Inquiries into Human Faculty and its Development. Londres: Dent. (Publicado originalmente en 1880.)

GARCÍA-MADRUGA, J. A. (1984), «Procesos de error en el razonamiento silogístico: doble procesamiento y estrategia de verificación». En Carretero, M., y Madruga, J. A. G. (eds.), Lecturas de psicologia del pensamiento. Madrid: Alianza.

GARCIA-MADRUGA, J. A.; MORENO, S.; CARRIEDO, N.; GUTIÉRREZ, F., y JOHNSON-LAIRD, P. N. (2001), «Are conjunctive inferences easier than disjunctive inferences? A comparison of rules and models», Quarterly Journal of Experimental Psychology, 54A, 613-632.

GÄRDENFORS, P. (1988), Knowledge in Flux. Cambridge, MA: MIT Press.

GÄRDENFORS, P. (1990), «The dynamics of belief systems: foundations vs. coherence theories», Revue Internationale de Philosophie, 172, 24-46.

GAREY, M., y JOHNSON, D. (1979), Computers and Intractability: a Guide to the Theory of NP- completeness San Francisco, CA: Freeman.

GARNHAM, A. (1987), Mental Models as Representations of Discourse and Text. Chichester: Ellis Horwood.

GARNHAM, A. (2001), Mental Models and the Representation of Anaphora. Hove, East Sussex: Psychology Press.

GARNHAM, A.; OAKHILL, J. V.; ERLICH, M-F., y CARREIRAS, M. (1995), «Representations and process in the interpretation of poronouns: new evidence from Spanish and French», Journal of Memory and Language, 34, 41-62.

GAUKER, C. (2005), Conditionals in Context. Cambridge, MA: MIT Press.

GAVANSKKI, I., y WELLS, G. L. (1989), «Counterfactual processing of normal

and exceptional events», Journal of Experimental Social Psychology, 25, 314-325.

GEERTZ, C. (1984), «Anti anti-relativism», American Anthropologist, 86, 263-277.

GELLNER, E. (1970), «Concepts and society». En Wilson, B. R. (ed.), Rationality. Oxford: Blackwell, pp. 18-49.

GELMAN, S. A., y MARKMAN, E. M. (1986), «Categories and induction in young children», Cognition, 23, 183-209.

GEMINIANI, G. C.; CARASSA, A., y BARA, E. G. (1996), «Causality by contact». En Oakhill, J., y Garnham, A. (eds.), Mental Models in Cognitive Science. Hove, East Sussex: Psychology Press, pp. 275-303.

GENTNER, D. (1983), «Structure mapping: a theoretical framework for analogy», Cognitive Science, 7, 155-170.

GENTZEN, G. (1969), «Investigations into logical deduction». En Szabo, M. E. (ed.), The Collected Papers of Gerhard Gentzen. Amsterdam: North-Holland, pp. 68-131. (Publicado originalmente en 1935.)

GEURTS, B. (2003), «Reasoning with quantifiers», Cognition, 86, 223-251.

GICK, M., y HOLYOAK, K. (1983), «Schema induction and analogical transfer», Cognitive Psychology, 15, 1-38.

GIERE, R. N. (1988), Explaining Science: a Cognitive Approach. Chicago, IL: University of Chicago Press.

GIERE, R. N. (ed.), (1992), Cognitive Models of Science. Minneapolis, MN: University of Minnesota Press.

GIGERENZER, G. (1996), «On narrow norms and vague heuristics: a reply to Kahneman and Tversky (1996)», Psychological Review, 103, 592-596.

GIGERENZER, G., y GOLDSTEIN, D. G. (1996), «Reasoning the fast and frugal way: models of bounded rationality», Psychological Review, 103, 650-669.

GIGERENZER, G., y HOFFRAGE, U. (1995), «How to improve bayesian reasoning without instruction: frequency format», Psychological Review, 102, 684-704.

GIGERENZER, G., y HUG, K. (1992), «Domain specific reasoning: Social contracts, cheating, and perspective change», Cognition, 43, 127-171.

GIGERENZER, G., y SELTEN, R. (eds.) (2002), Bounded Rationality: the Adaptive Toolbox. Cambridge, MA: MIT Press.

GIGERENZER, G.; TODD, P. M., y ABC RESEARCH GROUP (eds.) (1999), Simple Heuristics That Make Us Smart. Oxford: Oxford University Press.

GILBERT, D. T. (1989), «Thinking lightly about others: automatic components of the social inference process». En Uleman, J., y Bargh, J. A. (eds.), Unintended Thought. Nueva York: Guilford Press, pp. 189-211.

GILBERT, D. T., y WILSON, T. D. (2000), «Miswanting: Some problems in the forecasting of future affective states». En Forgas, J. (ed.), Thinking and Feeling: the Role of Affect in Social Cognition. Cambridge: Cambridge University Press, pp. 178-197.

GILBERT, D. T.; PINEL, E. G.; WILSON, T. C.; BLUMBERG, S. J., y WHEATLEY, T. P. (1998), «Immune neglect: a source of durability bias in affective forecasting», Journal of Personality and Social Psychology, 75, 617-638.

GILHOOLY, K. J.; LOGIC, R. H.; WETHERICK, N. E., y WYNN, V. (1993), «Working memory and strategies in syllogistic-reasoning tasks», Memory & Cognition, 21, 115-124.

GILHOOLY, K. J.; LOGIC, R. H., y WYNN, V. (1999), «Syllogistic reasoning tasks, working memory, and skill». European Journal of Cognitive Psychology, 11, 473-498.

GIMPEL, J. F. (1965), «A method of producing a Boolean funclion having an arbitrarily described prime implicant table», IEEE Transactions on Computing, 14, 484- 488.

GINSBERG, M. L. (1987), «Introduction». En Ginsberg, M. I. (ed.), Readings in

Nonmonotonic Reasoning. Los Altos, CA: Morgan Kaufinann, pp. 1-23.

GIROTTO, V., y GONZALEZ, M. (2001), «Solving probabilistic and statistical problems: a matter of question form and information structure», Cognition, 78, 247-276.

GIROTTO, V., y GONZALEZ, M. (2002), «Chances and frequencies in probabilistic reasoning: rejoinder to Hoffrage, Gigerenzer, Krauss and Martignon», Cognition, 84, 353-359.

GIROTTO, V., y GONZALEZ, M. (2006), «Norms and intuitions about chance». En Smith, L., y Voneche, J. (eds.). Norms and Development. Cambridge: Cambridge University Press, pp. 220-236.

GIROTTO, V., y JOHNSON-LAIRD, P. N. (2004), «The probability of conditionals», Psychologia, 47, 207-225.

GIROTTO, V.; LIGHT, P. H., y COLBOURN, C. J. (1988), «Pragmatic schemas and conditional reasoning in children», Quarterly Journal of Experimental Psychology, 40A, 342-357.

GIROTTO, V.; LEGRENZI, P., y RIZZO, A. (1991), «Event controllability in counterfactual thinking», Acta Psychologica, 78, 111-133.

GIROTTO, V.; MAZZOCCO, A., y TASSO, A. (1997), «The effect of premise order in conditional reasoning: a test of the mental model theory», Cognition, 63, 1-28.

GIROTTO, V.; JOHNSON-LAIRD, P. N.; LEGRENZI, P., y SONINO, M. (2000), «Reasoning to consistency: how people resolve logical inconsistencies». En García-Madruga, J. A.; Carriedo, N., y González- Labra, M. (eds.), Mental Models in Reasoning. Madrid: Universidad Nacional de Educación a Distancia, pp. 83-97.

GITTINGS, R. (ed.) (1970), Letters of John Keats. Oxford: Oxford University Press.

GLADWELL, M. (2005), Blink: the Power of Thinking without Thinking. Nueva York: Little, Brown and Co. [Ed. cast.: Blink: el poder de pensar sin pensar, Editorial Taurus, 2005.]

GLADWIN, T. (1970), East is a Big Bird. Cambridge, MA: Harvard University Press.

GLASGOW, J. I. (1993), «Representation of spatial models for geographic information systems». En Pissinou, N. (ed.), Proceedings of the ACM Workshop on Advances in Geographic Information Systems. Arlington, VA: Association for Computing Machinery, 112-117.

GLEICK, J. (1992), Genius: the Life and Science of Richard Feynman. Nueva York: Pantheon Books.

GLENBERG, A. M.; MEYER, M., y LINDEM, K. (1987), «Mental models contribute to foregroundng during text comprehension», Journal of Memory and Language, 26, 69-83.

GOEL, V., y DOLAN, R. J. (2004), «Differential involvement of left prefrontal cortex in inductive and deductive reasoning», Cognition, 93, B109-21.

GOETHE, J. W. von (1970), Theory of Colors. (Introducción de Judd, D. B.; trad.. Eastlake, C. L.). Cambridge, MA: MIT Press. (Traducción publicada inicialmente en 1840.)

GOFFMAN, E. (1959), The Presentation of Self in Everyday Life. Nueva York, NY: Doubleday.

GOLDING, W. (1955), The Inheritors. Londres: Faber and Faber. [Ed. cast.: Los herederos, trad. de L. Echávarri, Minotauro, 1993.]

GOLDVARG, Y., y JOHNSON-LAIRD, P. N. (2000), «Illusions in modal reasoning», Memory & Cognition, 28, 282-294.

GOLDVARG, Y., y JOHNSON-LAIRD, P. N. (2001), «Naïve causality: a mental model theory of causal meaning and reasoning. Cognitive Science, 25, 565-610.

GOODING, D. C. (2005), «Seeing the forest for the trees: Visualization, cognition, and scientific inference». En Gorman, M. E.; Tweney, R. D.; Gooding, D. C., y Kincannon, A. P. (eds.), Scientific and Technological Thinking. Mahwah, NJ: Erlbaum, pp. 173-217.

GOODMAN, N. (1972), «Seven strictures on similarity». En Goodman, N. (ed.),

Problems and Proyects. Nueva York: Bobbs-Merrill.

GOODMAN, N. (1983), Fact, Fiction, and forecast, 4ª ed. Cambridge, MA: Harvard University Press. (Publicado originalmente en 1954.)

GOODWIN, G., y JOHNSON-LAIRD, P. N. (2005a), «Reasoning about relations», Psychological Review, 112, 468-493.

GOODWIN, G., y JOHNSON-LAIRD, P. N. (2005b), «Reasoning about the relations between relations», Quarterly Journal of Experimental Pyschology, 59, 1-23.

GOODWIN, G., y JOHNSON-LAIRD, P. N. (2005c), «Illusions in transitive reasoning». En revisión.

GORDON, P. (2004), «Numerical cognition without words: evidence from Amazonia», Science, 306, 496-499.

GORMAN, M. E. (1992), Simulating Science: Heuristics, Mental Models, and Technoscientific Thinking. Bloomington, IN: Indiana University Press.

GORMAN, M. E.; TWENEY, R. D.; GOODING, D. C., y KINCANNON, A. P. (eds.) (2005), Scientific and Technological Thinking. Mahwah, NJ: Erlbaum.

GOSSE, E. (1923), Father and Son: a Study of Two Temperaments, 8ª ed. Nueva York: Scribner's Son. (Publicado originalmente en 1907.)

GOSSE, P. H. (1998), Omphalos: an Attempt to Untie the Geological Knot. Woodbridge, CT: Ox Bow Press. (Publicado originalmente en 1857.)

GOULD, J. L., y GOULD, C. G. (1994), The Animal Mind. Nueva York: Freeman.

GOULD, S. J. (1989), Wonderful Life: the Burgess Shale and the Nature of History. Nueva York: Penguin.

GOVIER, T. (1988), A Practical Study of Argument, 2ª ed. Belmont, CA: Wadsworth.

GRALINSKI, H. J., y KOPP, C. B. (1993), «Everyday rules for behavior:

mothers' requests to young children», Developmental Psychology, 29, 573-584.

GRANT, U. S. (1990), Memoirs and Selected Letters. Nueva York: Library of America. (Publicado originalmente en 1885.)

GREEN, D. W. (1995), «Externalization, counter-examples and the abstract selection task», Quarterly Journal of Experimental Psychology, 48A, 424-446.

GREEN, D. W. (1997), «Hypothetical thinking in the selection task: amplifying a model-based approach», Current Psychology of Cognition, 16, 93-102.

GREEN, D. W., y LARKING, R. (1995), «The locus of facilitation in the abstract selection task», Thinking & Reasoning, 1, 183-199.

GREEN, D. W., y MCCLOY, R. (2003), «Reaching a verdict», Thinking & Reasoning, 9, 307-333.

GREENE, E. J., y DARLEY, J. M. (1998), «Effects of necessary, sufficient, and indirect causation on judgments of criminal liability», Law and Human Behavior, 22, 429-451.

GREENE, J. D.; SOMMERVILLE, R. D.; NYSTROM, L. E.; DARLEY, J. M., y COHEN, J. D. (2001), «An fMRI investigation of emotional engagement in moral judgment», Science, 293, 2105-2108.

GREENE, S. (1992), «Multiple explanations for multiply quantified sentences: are multiple models necessary?», Psychological Review, 99, 184-187.

GREENE, T. R., y NOICE, H. (1988), «Influence of positive affect upon creative thinking and problem solving in children», Psychological Reports, 63, 895-898.

GREENWALD, A. G.; NOSEK, B. A., y BANAJI, M. R. (2003), «Understanding and using the Implicit Association Test: I. An improved scoring algorithm», Journal of Personality and Social Psychology, 85, 197-216.

GRICE, H. P. (1975), «Logic and conversation». En Davidson, D., y Harman, G. (eds.), The Logic of Grammar. Encino, CA: Dickenson, pp. 64-75. (Reimpreso en Grice, 1989.)

GRICE, H. P. (1989), Studies in the Ways of Words. Cambridge, MA: Harvard University Press.

GRIGGS, R. A. (1995), «The effects of rule clarification, decision justification, and selection instruction on Wason's abstract selection task». En Newstead, S. E., y Evans, J. St. B. T. (eds.), Perspectives on Thinking and Reasoning: Essays in Honor of Peter Wason. Hillsdale, NJ: Erlbaum, pp. 17-39.

GRUBER, H. E. (1994), «Insight and affect in the history of science». En Sternberg, R. J., y Davidson, J. (eds.), The Nature of Insight. Cambridge, MA: MIT Press, pp. 397-431.

GURDJIEFF, G. I. (1991), Life is Real Only Then, When "I Am". Nueva York: Viking Arkana.

GUYOTE, M. J., y STERNBERG, R. J. (1981), «A transitive-chain theory of syllogistic reasoning», Cognitive Psychology, 13, 461-525.

HACKING, I. (1975), The Emergence of Probability. Cambridge: Cambridge University Press. [Ed. cast.: El surgimiento de la probabilidad: un estudio filosófico de las ideas tempranas de la probabilidad, la inducción y la inferencia, trad. de J. Álvarez, Gedisa, 1995.]

HADAMARD, J. (1996), The Mathematician's Mind. Princeton, NJ: Universidad de Princeton Press. (Publicado originalmente como The Psychology of Invention in the Mathematical Field en 1945.)

HADJICHRISTIDIS, C.; STEVENSON, R. J.; OVER, D. E.; SLOMAN, S. A.; EVANS, J. St. B. T., y FEENEY, A. (2001), «On the evaluation of If p then q conditionals», Proceedings of the Twenty-third Annual Conference of the Cognitive Science Society. Hillsdale, NJ: Erlbaum, pp. 381-386.

HAGERT, G. (1984), «Modeling mental models: Experiments in cognitive modeling spatial relations». En O'Shea, T. (ed.), Advances in Artificial Intelligence. Amsterdam: North-Holland, pp. 179-188.

HALFORD, G. S. (1993), Children's Understanding: the Development of Mental Models. Hillsdale, NJ: Erlbaum.

HALFORD, G. S.; WILSON, W. H., y PHILLIPS, S. (1998), «Processing

capacity denned by relational complexity: Implications for comparative, developmental, and cognitive psychology», Behavioral and Brain Sciences, 21, 803-831.

HALLIDAY, S. (1999), The Great Stink of Londres: Sir Joseph Bazalgette and the Cleansing of the Victorian Capital. Stroud, Gloucestershire: Sutton.

HALMOS, P. (1957), «Nicolas Bourbaki», Scientific American, 196 (Mayo), 88-99.

HALPERN, J., y VARDI, M. Y. (1991), «Model checking vs. theorem proving: a manifesto». En Allen, J.; Fikes, R. E., y Sandewall, E. (eds.), Proceedings of the Second International Conference on Principles of Knowledge Representation and Reasoning, KR'91. San Mateo, CA: Morgan Kaufmann, pp. 325-334.

HAMMERTON, M. (1997), «Ignoring the base-rate -a robust phenomenon». Artículo presentado en la reunión de marzo de la Experimental Psychology Society, Universidad de Oxford, Inglaterra.

HAMPTON, J. A. (1979), «Polymorphous concept in semantic memory», Journal of Verbal Learning and Verbal Behavior, 18, 441-461.

HARDMAN, D. K. (1996), «Mental models: the revised theory brings new problems», Behavioral and Brain Sciences, 19, 542-543.

HARE, R. M. (1981), Moral Thinking: Its Levels, Method, and Point. Oxford Clarendon Press.

HARMAN, G. (1965), «The inference to the best explanation», Philosophical Review, 74, 88-95.

HARMAN, G. (1973), Thought. Princeton, NJ: Princeton University Press.

HARMAN, G. (1977), The Nature of Morality. Nueva York: Oxford University Press.

HARMAN, G. (1986), Change in View: Principles of Reasoning. Cambridge, MA: MIT Press, Bradford Book.

HARRE, R., y MADDEN, E. H. (1975), Causal Powers. Oxford: Blackwell.

HARRIS, P. L. (2000), The Work of the Imagination. Oxford: Blackwell.

HARRIS, P., y NÚÑEZ, M. (1996), «Understanding of permission rules by preschool children», Child Development, 67, 1572-1591.

HART, H. L. A., y HONORÉ, A. M. (1985), Causation in the Law, 2ª ed. Oxford: Clarendon Press. (Primera edición publicada en 1959.)

HASSON, U., y JOHNSON LAIRD, P. N. (2006), «How reasoning changes your belief in statements». En revisión.

HASTINGS, M. (2004), Armageddon: the Battle for Germany, 1944-1945. Nueva York: Knopf.

HAUSER, M.; CHOMSKY, N., y FITCH, W. T. (2002), «The language faculty: what is it, who has it, and how did it evolve?», Science, 298, 1569-1579.

HAYES-ROTH, F., y MCDERMOTT, J. (1978), «An interference matching technique for inducing abstractions», Communications of the Association for Computing Machinery, 21, 401-411.

HEGARTY, M. (1992), «Mental animation: Inferring motion from static diagrams of mechanical systems», Journal of Experimental Psychology: Learning, Memory, and Cognition, 18, 1084-1102.

HEGARTY, M. (2004), «Mechanical reasoning as mental simulation», Trends in Cognitive Sciences, 8, 280-285.

HEGARTY, M., y SIMS, V. K. (1994), «Individual differences in mental animation during mechanical reasoning», Memory & Cognition, 22, 411-430.

HEGARTY, M.; MAYER, S.; KRIZ, S., y KEEHNER, M. (2005), «The role of gestures in mental animation», Spatial Cognition and Computation, 5, 333-356.

HELLER, J. (1961), Catch-22, A Novel. Nueva York: Simon & Schuster. [Ed. cast.: Trampa 22, trad. de Flora Casas, RBA, 2005.]

HELMHOLTZ, H. (1962), Treatise on Physiological Optics. Nueva York: Dover. (Publicado originalmente en 1866.)

HENLE, M. (1962), «The relation between logic and thinking», Psychological Review, 69, 366-378.

HENLE, M. (1978), Prólogo a Revlin, R., y Mayer, R. E. (eds.), Human Reasoning. Washington, DC: Winston, pp. xiii-xviii.

HERINK, R. (ed.), (1980), The Psychotherapy Handbook: the A to Z Guide to More Than 250 Different Therapies in Use Today. Nueva York: New American Library.

HERSCHEL, J. F. W. (1987), A Preliminary Discourse on the Study of Natural Philosophy. Londres. Longman, Rees, Orme, Brown, y Green..., y John Taylor. Rep. Chicago: University of Chicago Press. (Publicado originalmente en 1831.)

HESSE, M. (1966), Models and Analogies in Science. Notre Dame, IN: Notre Dame University Press.

HEWITT, C. (1969), «PLANNER: a language for proving theorems in robots», International Joint Conference on Artificial Intelligence, 2, 295-302.

HILBERT, D. (1967), «The foundations of mathematics». En Van Heijenoort, J. (ed.), From Frege to Gödel: a Source Book in Mathematical Logic, 1879-1931. Cambridge, MA: Harvard University Press, pp. 464-479. (Publicado originalmente en alemán en 1928.)

HILPINEN, R. (ed.) (1971), Deontic Logic: Introductory and Systematic Readings. Dordrecht, Holanda: Reidel.

HILTON, D. J., y ERB, H.-P. (1996), «Mental models and causal explanation: judgements of probable cause and explanatory relevance», Thinking & Reasoning, 2, 273-308.

HINSLEY, F. H. (1975), British Intelligence in the Second World War: its Influence on Strategy and Operations. Londres: Her Majesty’s Stationary Office (HMSO).

HINSLEY, F. H., y STRIPP, A. (eds.) (1993), Code Breakers: the Inside Story of Bletchley Park. Oxford: Oxford University Press.

HINTIKKA, J. (1963), «The models of modality», Acta Philosophica Fennica,

16, 65-82.

HINTIKKA, J. (1973), Logic, Language-Games and Information: Kantian Themes in the Philosophy of Logic. Oxford: Clarendon Press.

HIRST, W., y WEIL, J. (1982), «Acquisition of the epistemic and deontic menaning of modals», Journal of Child Language, 9, 659-666.

HODGES, A. (2000), Alan Turing: the Enigma. Nueva York: Walker. (Primera edición publicada en 1992.)

HOGARTH, R. M., y EINHORN, H. J. (1992), «Order effects in belief updating: the belief-adjustment model», Cognitive Psychology, 24, 1-55.

HOLLAND, J. G., y SKINNER, B. F. (1961), The Analysis of Behavior: a Program for Self-instruction. Nueva York: McGraw-Hill.

HOLLAND, J. H.; HOLYOAK, K. J.; NISBETT, R. E., y THAGARD, P. R. (1986), Induction: Processes of Inference, Learning, and Discovery. Cambridge, MA: MIT Press.

HOLLARD, V., y DELIUS, J. D. (1982), «Rotational invariance in visual pattern recognition by pigeons and humans», Science, 218, 804-806.

HOLLON, S. D.; THASE, M. E., y MARKOWITZ, J. C. (2002), «Treatment and prevention of depression», Psychological Science in the Public Interest, 3, 1-39.

HOLYOAK, K. J., y GLASS, A. (1975), «The role of contradictions and counterexamples in the rejection of false sentences», Journal of Verbal Learning and Verbal Behavior, 14, 215-239.

HOLYOAK, K. J., y THAGARD, P. (1989), «Analogical mapping by constraint satisfaction», Cognitive Science, 13, 295-355.

HOOKE, R. (1705), The Posthumous Works of Robert Hooke, containing his Cutlerian Lectures, and Other Discourses. Londres: Richard Waller.

HOPCROFT, J. E., y ULLMAN, J. D. (1979), Formal Languages and Their Relation to Automata. Reading, MA: Addison-Wesley.

HORN, L. R. (1989), A Natural History of Negation. Stanford, CA: CSLI Publications.

HÖRNIG, R.; OBERAUER, K., y WEIDENFELD, A. (2005), «Two principles of premise integration in spatial reasoning», Memory & Cognition, 33, 131-139.

HOWSON, C., y URBACH, P. (1993), Scientific Reasoning: the Bayesian Approach, 2ª ed. Chicago, IL: Open Court.

HUGHES, G. E., y CRESSWELL, M. J. (1996), A New Introduction to Modal Logic. Londres: Routledge.

HULL, C. L. (1920), «Quantitative aspects of the evolution of concepts», Psychological Monographs, 28, número completo 123.

HULL, C. L. (1934), «The concept of the habit-family hierarchy and maze learning: Parte 1», Psychological Review, 41, 33-54.

HUME, D. (1978), A Treatise on Human Nature. Selby-Bigge, L. A. (ed.), 2 ª ed. Oxford: Oxford University Press. (Publicado originalmente en 1739.) [Ed. cast.: Tratado de la naturaleza humana, trad. de Félix Duque, Madrid, Tecnos, 2005.]

HUME, D. (1988), An Enquiry Concerning Human Understanding. Flew, A. (ed.), La Salle, IL: Open Court. (Publicado originalmente en 1748.)

HUMMEL, J. E., y HOLYOAK, K. J. (2001), «A process model of human transitive inference». En Gattis, M. (ed.), Spatial Schemas and Abstract Thought. Cambridge, MA: MIT Press, pp. 279-305.

HUMMEL, J. E., y HOLYOAK, K. J. (2003), «A symbolic-connectionist theory of relational inference and generalization», Psychological Review, 110, 220-264.

HUMPHREY, G. (1951), Thinking. Londres: Methuen.

HUMPHREY, N. (1996), Leaps of Faith: Science, Miracles, and the Search for Supernatural Consolation. Nueva York: Basic Books.

HUNT, E. B. (1999), «What is a theory of thought?». En Sternberg, R. J. (ed.), The Nature of Cognition. Cambridge, MA: MIT Press, pp. 3-49.

HUNT, E. B. (2002), Precis of Thoughts on Thought. Mahwah, NJ: Erlbaum.

HUNT, E. B.; MARIN, J., y STONE, P. T. (1966), Experiments in Induction. Nueva York: Academic Press.

HUNTER, I. M. L. (1957), «The solving of three terms series problems», British journal of Psychology, 48, 286-298.

HUTTENLOCHER, J. (1968), «Constructing spatial images: a strategy in reasoning», Psychological Review, 75, 550-560.

HUXLEY, T. H. (1893), «On the hypothesis that animals are automata, and its history». En Method and Results: Collected Essays Part One. Londres: Macmillan, pp. 199-250. (Publicado originalmente en 1874.)

INHELDER, B., y PIAGET, J. (1958), De la logique de l'enfant à la logique de l'adolescent. Paris: P.U.F. [Ed. cast.: De la lógica del niño a la lógica del adolescente, trad. de M. T. Cevasco, Buenos Aires, Paidós, 1972.]

INHELDER, B., y PIAGET, J. (1964), The Early Growth of Logic in the Child. Londres: Routledge & Kegan Paul.

INDER, R. (1987), «Computer simulation of syllogism solving using restricted mental models». Tesis doctoral sin publicar, Estudios cognitivos, Universidad de Edimburgo.

ISAAK, M. I., y JUST, M. A. (1995), «Constraints on thinking in insight and invention». En Sternberg, R. J., y Davidson, J. E. (eds.), The Nature of Insight. Cambridge MA: Bradford books, MIT Press, pp. 281-325.

ISEN, A. M.; DAUBMAN, K. A., y NOWICKI, G. P. (1987), «Positive affect facilitates creative problem solving», Journal of Personality and Social Psychology, 52, 1122-1131.

JACKENDOFF, R. (1996), «How language helps us think», Pragmatics & Cognition, 4, 1-34.

JACKSON R., y TARSITANO M. (1997), «Araneophagic jumping spiders discriminate between detour sites that do and do not lead to prey», Animal Behaviour, 53, 257-266.

JAKAB, P. L. (1990), Visions of a Flying Machine: the Wright Brothers and the Process of Invention. Washington, DC: Smithsonian Institution Press.

JAMES, W. (1884), «What is an emotion?», Mind, 9, 188-205.

JAMES, W. (1890), The Principles of Psychology, Vols. 1 y 2. Nueva York: Holt.

JAMES, W. (1907), Pragmatism— A New Name for Some Old Ways of Thinking. Nueva York: Longmans, Green & Co.

JAMES, W. (1956), The Will to Believe. Nueva York: Dover. (Publicado originalmente en 1896.)

JANG, K. L.; MCCRAE, R. R.; ANGLEITNER, A.; RIEMANN, R., y LIVESLEY, W. J. (1998), «Heritability of facet-level traits in a cross-cultural twin sample: support for a hierarchical model of personality», Journal of Personality and Social Psychology, 74, 1556-1565.

JANSSON, D. G., y SMITH, S. M. (1991), «Design fixation», Design Studies, 12, 3-11.

JANVEAU-BRENNAN, G., y MARKOVITS, H. (1999), «Reasoning with causal conditionals: developmental and individual differences», Developmental Psychology, 35, 904-911.

JEFFREY, R. (1981), Formal Logic: Its Scope and Limits, 2ª ed. Nueva York: McGraw-Hill.

JOHNSON, E. J., y TVERSKY, A. (1983), «Affect, generalization, and the perception of risk», Journal of Personality and Social Psychology, 45, 20-31.

JOHNSON, M. K. (1983), «A multiple-entry, modular memory system». En Bower, G. H. (ed.), The Psychology of learning and Motivation, Vol. 17. Nueva York: Academic Press, pp. 81-123.

JOHNSON, M. K., y REEDER, J. A. (1997), «Consciousness as meta-processing». En Cohen, J. D., y Schooler, J. W. (eds.), Scientific Approaches to Consciousness. Mahwah, NJ: Erlbaum, pp. 261-293.

JOHNSON, R. H., y BLAIR, J. A. (1994), Logical Self-Defense. Nueva York: McGraw-Hill.

JOHNSON-LAIRD, P. N. (1970a), «The perception and memory of sentences». En Lyons, J. (ed.), New Horizons in Linguistics. Harmondsworth, Middx: Penguin Books, pp. 261-270.

JOHNSON-LAIRD, P. N. (1970b), «The interpretation of quantified sentences». En Levelt, W. J. M., y Flores d'Arcais, G. B. (eds.), Advances in psicholinguistics. Amsterdam: North Holland Press, pp. 347-372.

JOHNSON-LAIRD, P. N. (1975), «Models of deduction». En Falmagne, R. J. (ed.), Reasoning: Representation and Process in Children and Adults. Hillsdale, NJ: Erlbaum, pp. 7-54.

JOHNSON-LAIRD, P. N. (1983), Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness. Cambridge, MA: Harvard University Press.

JOHNSON-LAIRD, P. N. (1993), Human and Machine Thinking. Hillsdale, NJ: Erlbaum.

JOHNSON-LAIRD, P. N. (1997), «Rules and illusions: a critical study of Rips's The Psychology of Proof». Minds and Machines, 7, 387-407.

JOHNSON-LAIRD, P. N. (1999), «Causation, mental models, and the law», Brooklyn Law Review, 65, 67-103.

JOHNSON-LAIRD, P. N. (2002a), «Peirce, logic diagrams, and the elementary operations of reasoning», Thinking & Reasoning, 8, 69-95.

JOHNSON-LAIRD, P. N. (2002b), «How jazz musicians improvise», Music Perception, 19, 415-442.

JOHNSON-LAIRD, P. N. (2005), «Flying bicycles: how the Wright brothers invented the airplane», Mind and Society, 27-48.

JOHNSON-LAIRD, P. N. (2006), How We Reason. Oxford: Oxford University Press.

JOHNSON-LAIRD, P. N., y BARA, B. G. (1984), «Syllogistic inference», Cognition, 16, 1-61.

JOHNSON-LAIRD, P. N., y BYRNE, R. M. J. (1989), Only reasoning. Journal of Memory and Language, 28, 313-330.

JOHNSON-LAIRD, P. N., y BYRNE, R. M. J. (1990), «Meta-logical problems: Knights, knaves, and Rips», Cognition, 36, 69-81.

JOHNSON-LAIRD, P. N., y BYRNE, R. M. J. (1991), Deduction. Hillsdale, NJ: Erlbaum.

JOHNSON-LAIRD, P. N., y BYRNE, R. M. J. (2002), «Conditionals: a theory of meaning, pragmatics, and inference», Psychological Review, 109, 646-678.

JOHNSON-LAIRD, P. N., y EHRLICH, K. (1982), «Spatial descriptions and referential continuity», Journal of Verbal Learning and Verbal Behavior, 21, 296-306.

JOHNSON-LAIRD, P. N., y HASSON, U. (2003), «Counterexamples in sentential reasoning», Memory & Cognition, 31, 1105-1113.

JOHNSON-LAIRD, P. N., y MANI, K. (1982), «The mental representation of spatial descriptions», Memory & Cognition, 10, 181-187.

JOHNSON-LAIRD, P. N., y OATLEY, K. (2000), «The cognitive and social construction of emotions». En Lewis, M., y Haviland, J. (eds.), Handbook of Emotions, 2ª ed. Nueva York: Guilford Press, pp. 458-475.

JOHNSON-LAIRD, P. N., y SAVARY, F. (1996), «Illusory inferences about probabilities», Acta Psychologica, 93, 69-90.

JOHNSON-LAIRD, P. N., y SAVARY, F. (1999), «Illusory inferences: a novel class of erroneous deductions», Cognition, 71, 191-229.

JOHNSON-LAIRD, P. N., y STEEDMAN, M. J. (1978), «The psychology of syllogisms», Cognitive Psychology, 10, 64-99.

JOHNSON-LAIRD, P. N., y TAGART, J. (1969), «How implication is understood», American Journal of Psychology, 82, 367-373.

JOHNSON-LAIRD, P. N., y WASON, P. C. (eds.) (1976), Thinking: Readings in Cognitive Science. Cambridge: Cambridge University Press.

JOHNSON-LAIRD, P. N.; LEGRENZI, P., y SONINO LEGRENZI, M. (1972), «Reasoning and a sense of reality», British Journal of Psychology, 63, 395-400.

JOHNSON-LAIRD, P. N.; OAKHILL, J. V., y BULL, D. (1986), «Children's syllogistic reasonig», Quarterly Journal of Experimental Psychology, 38A, 35-38.

JOHNSON-LAIRD, P. N.; BYRNE, R. M. J., y TABOSSI, P. (1989), «Reasoning by model: the case of multiple quantification», Psychological Review, 96, 658-673.

JOHNSON-LAIRD, P. N.; BYRNE, R. M. J., y SCHAEKEN, W. S. (1992), «Prepositional reasoning by model», Psychological Review, 99, 418-439.

JOHNSON-LAIRD, P. N.; LEGRENZI, P.; GIROTTO, V.; LEGRENZI, M., y CAVERNI, J.-P. (1999), «Naive probability: a mental model theory of extensional reasoning», Psychological Review, 106, 62-88.

JOHNSON-LAIRD, P. N.; LEGRENZI, P.; GIROTTO, P., y LEGRENZI, M. S. (2000), «Illusions in reasoning about consistency», Science, 288, 531-532.

JOHNSON-LAIRD, P. N.; GIROTTO, V., y LEGRENZI, P. (2004a), «Reasoning from inconsistency to consistency», Psychological Review, 111, 640-661.

JOHNSON-LAIRD, P. N.; LEGRENZI, P., y GIROTTO, V. (2004b), «How we detect logical inconsistencies», Current Directions in Psychological Science, 13, 41-45.

JOHNSON-LAIRD, P. N.; FROSCH, C., y GOLDVARG-STEINGOLD, G. (2006a), «Causes and enabling conditions: a reply to Kuhnmunch and Beller». En revisión.

JOHNSON-LAIRD, P. N.; MANCINI, F., y GANGEMI, A. (2006b), «A hyper emotion theory of psychological illnesses», Psychological Review, en prensa.

JOYCE, J. (1930), A Portrait of the Artist as a Young Man. Londres: Cape.

JUNG-BEEMAN, M.; BOWDEN, E. M.; HABERMAN, J.; FRYMIARE, J. L.; ARAMBEL-LIU, S.; GREENBLATT, R.; REBER, P. J., y KOUNIOS, J. (2004), «Neural activity when people solve verbal problems with insight», Public Library of Science Biology, 2(2), April. (Revista on-line.)

KAHANE, H. (1984), Logic and Contemporary Rhetoric, 4ª ed. Belmont, CA: Wadsworth.

KAHN, D. (1996), The Codebreakers. Nueva York: Scribner. (Primera edición publicada en 1967.)

KAHNEMAN, D. (1995), «Varieties of counterfactual thinking». En Roese, N. J., y Olson, J. M. (eds.), What Might Have Been: the Social Psychology of Counterfactual Thinking. Mahwah, NJ: Erlbaum, pp. 375-396.

KAHNEMAN, D., y FREDERICK, S. (2002), «Representativeness revisited: attribute substitution in intuitive judgement». En Gilovich, T.; Griffin, D., y Kahneman, D. (eds.), Heuristics and Biases: the Psychology of Intuitive Judgement. Cambridge: Cambridge University Press, pp. 49-81.

KAHNEMAN, D., y FREDERICK, S. (2005), «A model of heuristic judgment». En Holyoak, K. J., y Morrison, R. G. (eds.), The Cambridge Handbook of Thinking and Reasoning. Cambridge: Cambridge University Press, pp. 267-293. (Versión revisada de Kahneman y Frederick, 2002.)

KAHNEMAN, D., y MILLER, D. T. (1986), «Norm theory: comparing reality to its alternative», Psychological Review, 93, 75-88.

KAHNEMAN, D., y TVERSKY, A. (1972), «Subjective probability: a judgment of representativeness», Cognitive Psychology, 3, 430-454.

KAHNEMAN, D., y TVERSKY, A. (1973), «On the psychology of prediction», Psychological Review, 80, 237-251.

KAHNEMAN, D., y TVERSKY, A. (1979), «Prospect theory: an analysis of decision under risk», Econometrica, 47, 263-291.

KAHNEMAN, D., y TVERSKY, A. (1982), «The simulation heuristic». En Kahneman, D.; Slovic, P., y Tversky, A. (eds.), Judgment Under Uncertainty: Heuristics and Biases. Nueva York: Cambridge University Press, pp. 201-208.

KAHNEMAN, D., y TVERSKY, A. (1996), «On the reality of cognitive illusions: a reply to Gigerenzer's critique», Psychological Review, 103, 582-591.

KAHNEMAN, D.; SLOVIC, P., y TVERSKY, A. (eds.) (1982), Judgement Under Uncertainty. Cambridge: Cambridge University Press.

KAISER, M. K.; JONIDES, J., y ALEXANDER, J. (1986), «Intuitive reasoning on abstract and familiar physics probles», Memory & Cognition, 14, 308-312.

KALISH, C. (1998), «Reasons and causes: children's understanding of conformity to social rules and physical laws», Child Development, 69, 706-720.

KAMP, H. (1981), «A theory of truth and semantic representation». En Groenendijk, J. A. G.; Janssen, T. M. V., y Stokhof, M. B. J. (eds.), Fromal Methods in the Study of Language. Amsterdam: Mathematical Centre Tracts, pp. 277-322.

KAMP, H., y REYLE, U. (1993), From Discourse to Logic: Introduction to Modeltheroretic Semantics of Natural Language, Formal Logic and Discoruse Representation Theory. Dordrecht: Kluwer.

KANIGEL, R. (1991), The Man Who Knew Infinity: a Life of the Genius Ramanujan. Nueva York: Washington Square Press.

KANT, I. (1781), Kritik der reinen Vernunft. [Ed. cast.: Crítica de la razón pura, trad. de J. Rovira, Buenos Aires, Losada, 1960.]

KANT, I. (1974), Logic. (Publicado originalmente en1800.)

KANT, I. (1785), Groundwork of the Metaphysics of Morals. Cambridge: Cambridge University Press.

KAPLAN , C. A, y SIMON, H. A. (1990), «In search of insight», Cognitive Psychology, 22, 374-419.

KARMILOFF-SMITH, A. (1992), Beyond Modularity: a Developmental Perspective on Cognitive Science. Cambridge, MA: MIT Press.

KARMILOFF-SMITH, A., e INHELDER, B. (1974/75), «If you want to get ahead, get a theory», Cognition, 3, 195-212.

KATONA, G. (1940), Organizing and Memorizing: Studies in the Psychology of Learning and Teaching. Nueva York: Columbia University Press.

KEANE, M. T. (1988), Analogical Problem Solving. West Sussex: Ellis Horwood.

KEANE, M. T. (1989), «Modelling problem solving in Gestalt “insight” problems», Irish Journal of Psychology, 10, 201-215.

KEENE, G. B. (1992), The Foundations of Rational Argument. Lewinston, NY: Edward Mellen Press.

KEIL, F. C. (1989), Concepts, Kinds and Cognitive Development. Cambridge, MA: MIT Press.

KELLY, F. C. (ed.) (1996), Miracle at Kitti Hawk: the letters of Wilbur and Orville. Nueva York: Farrar, Straus & Young. (Publicado originalmente en 1951.)

KELVIN, Lord (1884), Notes of Lectures on Molecular Dynamics and the Wave Theory of Light. Baltimore: Johns Hopkins University.

KERMODE, F. (2003), Pieces of My Mind: Essays and Criticism 1958-2002. Nueva York: Farrar, Straus y Giroux.

KEYNES, H. (2002), Darwin, His Daughter and Human Evolution. Nueva York: Riverhead Books.

KEYNES, J. M. (1943), A Treatise on Probability. Londres: Macmillan.

KEYSAR, B.; LIN, S., y BARR, D. J. (2003), «Limits on theory of mind use in adults», Cognition, 89, 25-41.

KIHLSTROM, J. E.; BARNHARDT, T. M., y TATARYN, D. J. (1992), «Implicit perception». En Bornstein, R. E., y Pittman, T. S. (eds.), Perception Without Awareness. Nueva York: Guilford Press, pp. 17-54.

KLAHR, D., y SIMON, H. (1999), «Studies of scientific discovery: complementary approaches and convergent findings», Psychological Bulletin, 125, 524-543.

KLAUER, K. C.; JACOBSEN, T., y MIGULLA G. (1995), «Counterfactual processing: test of an hierarchical correspondence model», European Jhournal of Social Psychology, 25, 577-595.

KLAUER, K. C.; MUSCH, J., y NAUMER, B. (2000), «On belief bias in syllogistic reasoning», Psychological Review, 107, 852-884.

KLAYMAN, J., y HA, Y. W. (1987), «Confirmation, disconfirmation, and information in hypothesis testing», Psychological Review, 94, 211-228.

KLEER, J. de (1986), «An assumption-based TMS», Artificial Intelligence, 28, 127-162.

KLEIN, G. (1998), Sources of Power: How People make Decisions. Cambridge, MA: MIT Press.

KNAUFF, M., y JOHNSON-LAIRD, P. N. (2002), «Imagery can impede inference», Memory & Cognition, 30, 363-371.

KNAUFF, M., y MAY, E. (2006), «Mental imagery, reasoning, and blindness», Quarterly Journal of Experimental Psychology, 59, 161-177.

KNAUFF, M.; JOLA, C., y STRUBE, G. (2001), «Spatial resaoning: no need for visual information». En Montello, D. (ed.), Spatial Information Therory: Proceedings of COSIT'01. Nueva York: Springer-Verlag, pp. 447-457.

KNAUFF, M.; FANGMEIER, T.; RUFF, C. C., y JOHNSON-LAIRD, P. N. (2003), «Reaoning, models, and images: Behavioral measures and cortical activity», Journal of Cognitive Neuroscience, 4, 559-573.

KNEALE, W., y KNEALE, M. (1962), The Development of Logic. Oxford: Oxford University Press. [Ed. cast.: El desarrollo de la lógica, trad. de J. Muguerza, Tecnos, 1972.]

KNIGHT, R. T., y GRABOWECKY, M. (1995), «Escape from linear time: prefrontal cortex and conscious experience». En Gazzinaga, M. S. (ed.), The Cognitive Neurosciences. Cambridge, MA: MIT Press, pp. 1357-1556.

KNOBLICH, G.; OHLSSON, S.; HAIDER, H., y RHENIUS, D. (1999), «Constraint relaxation and chunk decomposition in insight problem-solving»,

Journal of Experimental Psychology: Learning, Memory & Cognition, 25, 1534-1556.

KOHLBERG, L. (1986), «A current statement of some theroretical issues». En Modgil, S., y Modgil, C. (eds.), Lawrence Kohlberg: Consensus and Controversy. Filadelfia: Falmer, pp. 485-546.

KÖHLER, W. (1947), Gestalt Psychology. Nueva York: Liveright. (Publicado originalmente en 1929.)

KOLODNER, J. (1993), Case-Based Reasoning. San Mateo, CA: Morgan Kaufman.

KORF, R. E. (1980), «Toward a model of representational changes», Artificial Intelligence, 14, 41-78.

KOSLOWSKI, B. (1996), Theory and Evidence: the Development of Scientific Reasoning. Cambridge, MA: MIT Press.

KOSSLYN, S. M. (1980), Image and Mind. Cambridge, MA: Harvard University Press.

KOSSLYN, S. M.; BALL, T. M., y REISER, B. J. (1978), «Visual images preserve metric spatial information: evidence from studies of image scanning», Journal of Experimental Psychology: Human Perception and Performance, 4. 47-60.

KOZACZUK, W., y STRASZAK, J. (2004), Enigma: How the Poles Broke the Enigma Code. Nueva York: Hippocrene Books.

KRIPKE, S. (1963), «Semantical considerations on modal logic», Acta Philosophica Fennica, 16, 83-94.

KROEBER, T. (1961), Ishi in Two Worlds: a Biography of the Last Wild Indian in North America. Berkeley, CA: University of California Press.

KUBOVY, M. (1986), Psychologyof Perspective and Renaissance Art. Cambridge University Press.

KUHN, D. (1977), «Conditional reasoning in children», Developmental

Psychology, 13, 342-353.

KUHN, D. (1991), The Skills of Argument. Cambridge: Cambridge University Press.

KUHN, T. (1970), The Structure of Scientific Revolutions, 2ª ed. Chicago, IL: University of Chicago Press.

KUHNMÜNCH, G., y BELLER, S. (2005), «Causes and enabling conditions-through mental models of linguistic cues?», Cognitive Science, 29, 1077-1090.

KULKARNI, D., y SIMON, H. A. (1988), «The process of scientific discovery: the strategy of experimentation», Cognitive Science, 12, 139-175.

KYLLONEN, P., y CHRISTAL, R. (1990), «Reasoning ability is (little more than), working-memory capacity?», Intelligence, 14, 389- 33.

LANDAU, B.; SPELKE, E., y GLEITMAN, H. (1984), «Spatial knowledge in a young blind child», Cognition, 16, 225-260.

LANDAUER, T. K. (1986), «How much do people remember? Some estimates of the quantity of learned information in long-term memory», Cognitive Science, 10, 477-493.

LANGLEY, P.; SIMON, H. A.; BRADSHAW, G. L., y ZYTKOW, J. M. (1987), Scientific Discovery. Cambridge, MA: MIT Press.

LAPLACE, P. S. de (1951), Philosophical Essay on Probabilities. Nueva York: Dover. (Publicado originalmente en 1820.)

LARKIN, J., y SIMON, H. (1987), «Why a diagram is (sometimes), worth 10,000 words», Cognitive Science, 11, 65-99.

LASHLEY, K. S. (1958), «Cerebral organization and behavior». En Solomon, H. C.; Cobb, S., y Penfield, W. (eds.), The Brain and Human Behavior, Vol. 36. Association for Research in Nervous and Mental Disorders, Research Publications. Baltimore: Williams y Wilkin, pp. 1-18.

LAZARUS, R. S. (1982), «Thoughts on the relationship between emotion and cognition», American Psychologist, 37, 1019-1024.

LEAHY, R. L. (ed.) (2004), Contemporary Cognitive Therapy: Theory, Research, and Practice. Londres: Guilford Press.

LEDOUX, J. (1996), The Emotional Brain: the Mysterious Underpinnings of Emotional Life. New York: Simon & Schuster.

LEE, N. Y. L., y JOHNSON-LAIRD, P. N. (2004), «Creative strategies in problem solving», Proceedings of the Twenty-Sixth Annual Conference of the Cognitive Science Society. Mahwah, NJ: Erlbaum, pp. 813-818.

LEE, N. Y. L., y JOHNSON-LAIRD, P. N. (2005), «Synthetic reasoning and the reverse engineering of Boolean circuits», Proceedings of the Twenty-Seventh Annual Conference of the Cognitive Science Society. Mahwah, NJ: Erlbaum, pp. 1260-1265.

LEE, N. Y. L.; GOODWIN, G. P., y JOHNSON-LAIRD, P. N. (2006), «The psychology of Su Doku». En revisión.

LEGRENZI, P.; GIROTTO, V., y JOHNSON-LAIRD, P. N. (2003), «Models of consistency», Psychological Science, 14, 131-137.

LEGRENZI, M. S.; LEGRENZI, P.; GIROTTO, V., y JOHNSON-LAIRD, P. N. (2001-2), «The science of finding», Bulletin of People-Environment Studies, 20, 7-9.

LEIBNIZ, G. (1765), New essays concerning Human Understanding, [Ed. cast.: Nuevo Tratado sobre el entendimiento humano, trad. de J. Echevarría Ezponda. Madrid. Editora Nacional, 1977.]

LEON, G. R. (1990), Case Histories of Psychopathology. Cuarta Ediciòn. Boston: Allyn & Bacon.

LEPPER, M. R.; ROSS, L., y LAU, R. R. (1986), «Persistence of inaccurate beliefs ab out the self: perseverance effects in the classroom», Journal of Personality and Social Psychology, 50, 482-491.

LESLIE, A. M. (1984), «Spatiolemporal contiguity and perception of causality in infants», Perception, 13, 287-305.

LESLIE, A. M. (1987), «Pretense and representation: the origins of “Theory of

Mind”», Psychological Review, 94, 412-426.

LESLIE, A. M. (1990), «Pretense, autism and the basis of “Theory of mind”», The Psychologist, 3, 120-123.

LESLIE, A. M., y THAISS, L. (1992), «Domain specificity in conceptual development: neuropsychological evidence from autism», Cognition, 43, 225-251.

LEVESQUE, H. J. (1986), «Making believers out of computers», Artificial Intelligence, 30, 81-108.

LEVI, I. (1991), The Fixation of Belief and its Undoing: Changing Beliefs through Inquiry. Nueva York, NY: Cambridge University Press.

LEVI, I. (1996), «For the Sake of Argument: Ramsey test conditionals», Inductive Inference and Nonmonotonic Reasoning. Cambridge: Cambridge University Press.

LEVINSON, S. C. (1996), «Frames of reference and Molyneaux’s question: crosslinguistic evidence». En Bloom, P.; Peterson, M. A.; Nadel, L., y Garrett, M. F. (eds.), Language and Space. Cambridge, MA: MIT Press, pp. 109-169.

LEWIS, C. (1986), «A model of mental model construction», Proceedings of CHI ‘86 Conference on Human Factors in Computer Systems. Nueva York: Association for Computing Machinery, pp. 306-313.

LEWIS, D. K. (1973), Counterfactuals. Oxford: Basil Blackwell.

LEWONTIN, R. C. (1990), «The evolution of cognition». En Osherson, D. N., y Smith, E. E. (eds.), An Invitation to Cognitive Science, Vol. 3. Thinking. Cambridge, MA: MIT Press. pp. 229-246.

LI, M., y VITANYI, P. M. B. (1993), An Introduction to Kolmogorov Complexity and Its Applications. Berlin: Springer-Verlag.

LIBERMAN, N., y KLAR, Y. (1996), «Hypothesis testing in Wason’s selection task: social exchange cheating detection or task understanding», Cognition, 58, 127-156.

LICHTENBERG, G. C. (1958), Aphorismen. Zurich: Manesse.

LIN, Y. (1937), The Importance of Living. Nueva York: Reynal & Hitchcock.

LIPTON, P. (1991), Inference to the Best Explanation. Londres: Routledge.

LITTLEWOOD, J. E. (1953), A Mathematician's Miscellany. Londres: Meuthen.

LIVESLEY, W. J.; JANG, K. L., y VERNON, P. A. (1993), «Genetic and environmental contributions to dimensions of personality disorder», American Journal of Psychiatry, 150, 12-20.

LOBER, K., y SHANKS, D. R. (2000), «Is causal induction based on causal power? Critique of Cheng (1997)», Psychological Review, 107, 195-212.

LOCKE, J. (1979), An Essay Concerning Human Understanding. Nidditch, P. H. (ed.), Oxford: Oxford University Press. (Publicado originalmente en 1689.)

LONGFORD, E. (1969), Wellington: the Years of the Sword. Nueva York: Harper and Row.

LÓPEZ, F. J.; COBOS, P. L.; CANO, A., y SHANKS, D. R. (1998), «The rational analysis of human causal and probability judgment». En Oaksford, M., y Chater, N. (eds.), Rational Models of Cognition. Oxford: Oxford University Press, pp. 314-352.

LOVE, R., y KESSLER, C. (1995), «Focussing in Wason's selection task: content and instruction effects», Thinking & Reasoning, 1, 153-182.

LOWENSTEIN, R. (2000), When Genius Failed: the Rise and Fall of Long-Term Capital Mangagement. Nueva York: Random House.

LUCHINS, A. S., y LUCHINS, E. S. (1950), «New experimental attempts at preventing mechanization in problem solving», Journal of General Psychology, 42, 279-297.

LURIA, A. R. (1934), «The second psychological expedition lo central Asia», Journal of Genetic Psychology, 41, 255-259.

LURIA, A. R. (1976), Cognitive Development: Its Cultural and Social

Foundations. Cambridge, MA: Harvard University Press.

LYCAN, W. G. (1991), «MPP, RIP». Artículo sin publicar. Universidad de Carolina del Norte.

LYCAN, W. G. (2001), Real Conditionals. Oxford: Clarendon Press.

MACGREGOR, J. N.; ORMEROD, T. C., y CHRONICLE, E. P. (2001), «Information-processing and insight: a process model of performance on the nine-dot and related problems», Journal of Experimental Psychology: Learning, Memory and Cognition, 27, 176-201.

MACKIE, J. L. (1980), The Cement of the Universe: a Study in Causation, 2ª ed. Oxford: Oxford University Press.

MACNAMARA, J. (1986), A Border Dispute: the Place of Logic in Psychology. Cambridge, MA: Bradford Books, MIT Press.

MAGNANI, L., y NERSESSIAN, N. (eds.) (2002), Model-Based Reasoning: Science, Technology, Values. Nueva York: Kluwer Academic/Plenum.

MAGNANI, L.; NERSESSIAN, N., y THAGARD, P. (eds.) (1999), Model-based Reasoning in Scientific Discovery. Nueva York: Kluwer Academic.

MAGUIRE, E. A.; GADIAN, D. G.; JOHNSRUDE, I. S.; GOOD, C. D.; ASHBURNER, J.; FRACKOWIAK, R. S. J., y FRITH, C. D. (2000), «Navigation-related structural change in the hippocampi of taxi drivers», Proceedings of the National Academy of Science USA, 97, 4398-4403.

MAIER, N. R. F. (1931), «Reasoning in humans: II. The solution of a problem and its appearance in consciousness», Journal of Comparative Psychology, 12, 181-194.

MALLESON, A. (1973), Need Your Doctor be so Useless. Londres: Allen & Unwin.

MALTHUS, T. R. (2004), An Essay on the Principle of Population. Appleman, P. (ed.), 2ª ed. Nueva York: Norton. (Publicado originalmente en 1798.) [Ed. cast.: Ensayo sobre el principio de la población, trad. de J. A. Moral Santín, Akal, 1990.]

MANDLER, J. (2004), «Thought before language», Trends in Cognitive Sciences, 8, 508-513.

MANKTELOW, K. I. (1999), Reasoning and Thinking. Hove, East Sussex: Psychology Press.

MANKTELOW, K. I. (2004), «Reasoning and rationality: the pure and the practical». En Manktelow, K. I., y Chung, M. C. (eds.), Psychology of Reasoning: Theoretical and Historical Perspectives. Nueva York: Psychology Press, pp. 157-177.

MARCUS, G. F. (2001), The Algebraic Mind: Integrating Connectionism and Cognitive Science. Cambridge, MA: MIT Press.

MARCUS, G. F.; VIJAYAN, S.; BANDI RAO, S., y VISHTON, P. M. (1999), «Rule learning by seven- month-old infants», Science, 283, 77-80.

MARKMAN, A., y DIETRICH, E. (2000), «Extending the classical view of representation», Cognitive Psychology, 40, 138-171.

MARKMAN, E. M. (1989), Categorization and Naming in Young Children: Problems in Induction. Cambridge, MA: MIT Press.

MARKOVITS, H. (1993), «The development of conditional reasoning: a Piagetian reformulation of mental models theory», Merrill-Palmer Quarterly, 39, 133-160.

MARKOVITS, H., y BARROUILLET, P. (2002), «The development of conditional reasoning: a mental model account», Developmental Review, 22, 5-36.

MARKOVITS H., y HANDLEY, S. J. (2005), «Is inferential rasoning just probabilistic reasoning in disguise?», Memory & Cognition, 33, 1315-1323.

MARKOVITS, H., y QUINN, S. (2002), «Eficiency of retrieval correlates with “logical” reasoning from causal conditional premises», Memory & Cognition, 30, 696-706.

MARKOVITS, H., y VACHON, R. (1990), «Conditional reasoning, representation, and level of abstraction», Developmental Psychology, 26, 942-

951.

MARKOVITS, H.; FLEURY, M.-L.; QUINN, S., y VENET, M. (1998), «The development of conditional reasoning and the structure of semantic memory», Child Development, 64, 742-755.

MARKOVITS, H.; DOYON, C, y SIMONEAU, M. (2002), «Individual differences in working memory and conditional reasoning with concrete and abstract content», Thinking & Reasoning, 8, 97-107.

MARR, D. (1982), Vision: a Computational Investigation into the Human Representation and Processing of Visual Information. San Francisco, CA: W. H. Freeman.

MARSH, A. A.; ELFENBEIN, H. A., y AMBADY, N. (2003), «Nonverbal "accents": Cultural differences in facial expressions of emotion», Psychological Science, 14, 373-376.

MARTIN-CORDERO, J., y GONZALEZ-LABRA, M. J. (1994), «Amnesic mental models do not completely spill the beans of deductive reasoning», Behavioral and Brain Sciences, 17, 773-774.

MASLOW, A. H. (1938), Cases in Personality and Abnormal Psychology. Nueva York: Brooklyn College Press.

MATHEWS, A.; MACKINTOSH, B., y FULCHER, E. P. (1997), «Cognitive biases in anxiety and attention to threat», Trends in Cognitive Sciences, 1, 340-345.

MAXWELL, J. C. (1911), «Diagram», The Encyclopaedia Britannica, Vol. XVIII. Nueva York: the Encylopaedia Britannica Company.

MAY, E. R., y ZELIKOW, P. D. (eds.) (2001), The Kennedy Tapes: Inside the Whitehouse during the Cuban Missile Crisis. Concise ed. Nueva York: Norton.

MAYR, E. (1982), The Growth of Biological Thought: Diversity, Evolution, and Inheritance. Cambridge, MA: Belknap, Harvard University Press.

MAZZOCCO, A.; LEGRENZI, P., y RONCATO, S. (1974), «Syllogistic inference: the failure of the atmosphere effect and the conversion hypothesis»,

Italian Journal of Psychology, 2, 157-172.

MCARTHUR, L. (1972), «The how and what of why: some determinants and consequences of causal attribution», Journal of Personality and Social Psychology, 22, 171-193.

MCCLELLAND, J. L., y RUMELHART, D. E. (1986), Parallel Distributed Processing. Explorations in the Microstructure of Cognition, Vol. 2: Psychological and Biological Models. Cambridge, MA: MIT.

MCCLOSKEY, M. (1983), «Naive theories of motion». En Gentner, D., y Stevens, A. L. (eds.), Mental Models. Hillsdale, NJ: Erlbaum, pp. 299-324.

MCCLUSKEY, E. J. (1956), «Minimization of Boolean functions», Bell Systems Technical Journal, 35, 1417-1444.

MCFARLAND, M. W. (ed.) (2001), The Papers of Wilbur and Orville Wright, 2 vols. Nueva York: McGraw-Hill. (1ª ed., 1953.)

MCFARLANE, G. (1984), Alexander Fleming: the Man and the Myth. Oxford: Oxford University Press.

MCPECK, J. E. (1981), Critical Thinking and Education. Oxford: Robertson.

MEDIN, D. L.; GOLDSTONE, R. L., y CENTNER, D. (1993), «Respects for similarity», Psychological Review, 100, 254-278.

MEDVEC, V. H.; MADEY, S. F., y GILOVICH, T. (1995), «When less is more: counterfactual thinking and satisfaction among Olympic medalists», Journal of Personality and Social Psychology, 69, 603-610.

MEDVEDEV, Z. A. (1990), The Legacy of Chernobyl. Now York: W.W. Norton.

MEHLER, J.; JUSCZYK, P.; LAMBERTZ, G.; HALSTED, N.; BERTONCINI, J., y AMICL-TISON, C. (1988), «Aprecursor of language acquisition in young infants», Cognition, 29, 143-178.

MELLERS, B. A., y MCGRAW, P. (2001), «Anticipated emotions as guides to choice», Current Directions in Psychological Science, 6, 210-214.

MELLERS, B. A.; SCHWARTZ, A., y COOKE, A. (1998), «Decision making and valuation», Annual Review of Psychology, 49, 447-477.

METCALFE, J., y WEIBE, D. (1987), «Intuition in insight and noninsight problem solving», Memory & Cognition, 15, 238-246.

METZLER, J., y SHEPARD, R. N. (1982), «Transformational studies of the internal representations of three-dimensional objects». En Shepard, R. N., y Cooper, L. A. (eds.), Mental Images and Their Transformations. Cambridge, MA: MIT Press, pp. 25-71. (Publicado originalmente en Solso, R. L. (ed.), Theories in Cognitive Psychology: the Loyola Symposium. Hillsdale, NJ: Erlbaum, 1974.)

MICHALSKI, R. S. (1983), «A theory and methodology of inductive learning». En Michalski, R. S.; Carbonell, J. G., y Mitchell, T. M. (eds.), Machine Learning: an Artificial Intelligence Approach. Los Altos, CA: Morgan Kaufmann, pp. 83-134.

MICHOTTE, A. (1963), The Perception of Causality. Londres: Methuen. (Publicado inicialmente en 1946.)

Mill, J. S. (1874), A System of Logic, Ratiocinative and Inductive: Being a Connected View of the Principles of Evidence and the Methods of Scientific Evidence, 8ª ed. Nueva York: Harper. (Primera edición publicada en 1843.)

MILL, J. S. (1874), A System of Logic, Ratiocinative and Inductive: Being a Connected View of the Principles of Evidence and the Methods of Scientific Evidence, 8ª edición, Nueva York: Harper. (Primera edición publicada en 1843.)

MILL, J. S. (1998), Utilitarianism. Oxford: Oxford University Press. (Publicado originalmente en 1863.)

MILLER, A. (1984), Imagery in Scientific Thought: Creating 20th-century Physics. Boston, MA: Birkhauser.

MILLER, G. A. (1968), The Psychology of Communication: Seven Essays. Harmondsworth, Middlesex: Penguin.

MILLER, G. A., y JOHNSON-LAIRD, P. N. (1976), Language and Perception. Cambridge, MA: Harvard University Press. Cambridge: Cambridge University

Press.

MILLER, G. A.; GALANTER, E., y PRIBRAM, K. H. (1960), Plans and the Structure of behavior. NewYork: Henry Holt. [Ed. cast.: Planes y estructura de la conducta, trad. de R. Fernández González, Debate, 1983.]

MINSKY, M. (1968), «Matter, mind, and models». En Minsky, M. (ed.), Semantic Information Processing. Cambridge, MA: MIT Press, pp. 425-432.

MINSKY, M. (1975), «Frame-system theory». En Schank, R. C., y Webber, B. L. (eds.), Theoretical Issues in Natural Language Processing. Reimpreso en Johnson-Laird, P. C., y Wason, P. C. (eds.), Thinking: Readings in Cognitive Science. Cambridge: Cambridge University Press, 1977, pp. 355-376.

MINSKY, M. (1985), The Society of Mind. Nueva York: Simon y Schuster.

MINSKY, M. L., y PAPERT, S. A. (1969), Perceptrons. Cambridge, MA: MIT Press.

MITCHELL, M. (1998), An Introduction to Genetic Algorithms. Cambridge, MA: MIT Press.

MIYAKE, N. (1986), «Constructive interaction and the iterative process of understanding», Cognitive Science, 10, 151-177.

MIYAMOTO, J. M.; GONZÁLEZ, R., y TU, S. (1995), «Compositional anomalies in the semantics of evidence». En Busemeyer, J.; Hastie, R., y Medin, D. (eds.), Decision Making from a Cognitive Perspective. (Vol. 32. The Psychology of Learning and Motivation.) Nueva York: Academic Press, pp. 319-383.

MONTAGUE, R. (1974), Formal Philosophy: Selected Papers. New Have: Yale University Press.

MORRIS, A. K., y SLOUTSKY, V. M. (1998), «Understanding of logical necessity: developmental antecedents ang cognitive consequences», Child Development, 69, 721-741.

MORRIS, B. J., y SLOUTSKY, V. M. (2002), «Children’s solutions of logical versus empirical problems: what’s missing and what develops», Cognitive

Development 16, 907-928.

MORRIS, M. W.; NISBETT, R. E., y PENG, K. (1995), «Causal attribution across domains and cultures». En Sperber, D.; Premack, D., y Premack, A. J. (eds.), Causal Cognition: a Multidisciplinary Debate. Nueva York: Oxford University Press, pp. 577-614.

MOSTOWSKI, A. (1957), «On a generalization of quantifiers», Fundamenta Mathematicae, 44, 12-36.

MOXEY, L. M., y SANFORD, A. J. (1993), Communicating Quantities: a Psychological Perspective. Hove, Sussex: Erlbaum.

MOXEY, L. M., y SANFORD, A. J. (2000), «Communicating quantities: a review of psycholinguistic evidence of how expressions determine perspectives», Applied Cognitive Psychology, 14, 237-255.

MURPHY, G. L., y MEDIN, D. L. (1985), «The role of theories in conceptual coherence», Psychological Review, 92, 289-316.

MURRAY, G. (1915), The Stoic Philosophy. Nueva York: G. P. Putnam.

NEISSER, U. (1963), «The imitation of man by machine», Science, 139, 193-197.

NEISSER, U., y WEENE, P. (1962), «Hierarchies in concept attainment», Journal of Experimental Psychology, 64, 640-645.

NESSE, R. M. (1991), «What good is feeling bad?», The Sciences, Noviembre/Diciembre, 30-37.

NETH, H., y JOHNSON-LAIRD, P. N. (1999), «The search for counterexamples in human reasoning», Proceedings of the Twenty First Annual Conference of the Cognitive Science Society, 806.

NEWELL, A. (1980), «Physical symbol systems», Cognitive Science, 2, 135-184.

NEWELL, A. (1981), «Reasoning, problem solving and decision processes: the problem space as a fundamental category». En Nickerson, R. (ed.), Attention and

Performance, Vol. 8. Hillsdale, NJ: Erlbaum, pp. 693-718.

NEWELL, A. (1990), Unified Theories of Cognition. Cambridge. MA: Harvard Press.

NEWELL, A., y SIMON, H. A. (1972), Human Problem Solving. Englewood Cliffs, NJ: Prentice Hall.

NEWSOME, M. R., y JOHNSON-LAIRD, P. N. (1996), «An antidote to illusory inferences?». En Cottrell, G. W. (ed.), Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society. Mahwah, NJ: Erlbaum, p. 820.

NEWSTEAD, S. E., y GRIGGS, R. A. (1983), «The language and thought of disjunction». En Evans, J. St. B. T. (ed.), Thinking and Reasoning: Psychological Approaches. Londres: Routledge & Kegan Paul, pp. 76-106.

NEWSTEAD, S. E.; POLLARD, P., y GRIGGS, R. A. (1986), «Response bias in relational reasoning», Bulletin of the Psychonomic Society, 24, 95-98.

NEWSTEAD, S. E.; ELLIS, M. C.; EVANS, J. ST. B. T., y DENNIS, I. (1997), «Conditional reasoning with realistic material», Thinking & Reasoning, 3, 49-76.

NEWSTEAD, S. E.; HANDLEY, S. J., y BUCK, E. (1999), «Falsifying mental models: testing the predictions of theories of syllogistic reasoning», Memory & Cognition, 27, 344-354.

NEWSTEAD, S. E.; THOMPSON, V. A., y HANDLEY, S. J. (2002), «Generating alternatives: a key component of human reasoning?», Memory & Cognition, 30, 129-137.

NEWTON, I. (1952), Opticks: a Treatise of the Reflections, Refractions, Inflections & Colours of Light. Nueva York: Dover. (Basado en la 4ª ed., Londres, 1730.)

NICHOLL, C. (2004), Leonardo da Vinci: Flights of the Mind. Nueva York: Viking.

NICHOLS, S. (2002), «Norms with feeling: towards a psychological account of moral judgment», Cognition, 84, 221-236.

NICKERSON, R. S. (1996), «Hempel's paradox and Wason's selection task: logical and psychological problems of confirmation», Thinking & Reasoning, 2, 1-31.

NIEDENTHAL, P. M.; TANGNEY, J. P., y GAVANSKI, I. (1994), «"If only I weren't" versus "if only I hadn't": Distinguishing shame and guilt in counterfactual thinking», Journal of Personality and Social Psychology, 67, 585-595.

NISBETT, R. E. (2003), The Geography of Thought: How Asians and Westerners Think Differently ... and Why. Nueva York: the Free Press.

NISBETT, R. E., y WILSON, T. (1977), «Telling more than we can know: verbal reports on mental processes», Psychological Review, 84, 231-259.

NISBETT, R. E.; KRANTZ, D. H.; JEPSON, D., y KUNDA, Z. (1983), «The use of statistical heuristics in everyday inductive reasoning», Psychological Review, 90, 339-363.

NORTHCUTT, W. (2000), The Darwin Awards: Evolution in Action. Nueva York: Penguin Putnam.

NORVIG, P. (1992), Paradigms of Artificial Intelligence Programming: Case Studies in Common Lisp. San Mateo, CA: Morgan Kaufman.

NORENZAYAN, A.; SMITH, E. E.; JUN KIM, B., y NISBETT, R. E. (2002), «Cultural preferences for formal versus intuitive reasoning», Cognitive Science, 26, 653-684.

NOVECK, I. A. (2001), «When children are more logical than adults: experimental investigations of scalar implicature», Cognition, 78, 165-188.

NUNDY, S. (1995), «The effects of emotion on human inference: towards a computational model». Tesis doctoral sin publicar, Graduate Department of Education, Universidad de Toronto.

OAKHILL, J. V., y JOHNSON-LAIRD, P. N. (1985), «The effects of belief on the spontaneous production of syllogistic conclusions», Quarterly Journal of Experimental Psychology, 37A, 553-569.

OAKHILL, J.; JOHNSON-LAIRD, P. N., y GARNHAM, A. (1989), «Believability and syllogistic reasoning», Cognition, 31, 117-140.

OAKHILL, J.; GARNHAM, A., y JOHNSON-LAIRD, P. N. (1990), «Belief bias in syllogistic reasoning». En Gilhooly, K.; Keane, M.; Logic, R., y Erdos, G. (eds.), Lines of Thinking, Vol. I. Nueva York: Wiley, pp. 125-138.

OAKSFORD, M., y CHATER, N. (1991), «Against logicist cognitive science», Mind & Language, 6, 1-38.

OAKSFORD, M., y CHATER, N. (1994), «A rational analysis of the selection task as optimal data selection», Psychological Review, 101, 608-631.

OAKSFORD, M., y CHATER, N. (1998), Rationality in an Uncertain World. Hove, UK: Psychology Press.

OAKSFORD, M., y CHATER, N. (2001), «The probabilistic approach to human reasoning», Trends in Cognitive Sciences, 5, 349-357.

OAKSFORD, M., y STENNING, K. (1992), «Reasoning with conditionals containing negated constituents», Journal of Experimental Psychology: Learning, Memory, and Cognition, 18, 835-854.

OAKSFORD, M.; MORRIS, E.; GRAINGER, R., y WILLIAMS, J. M. G. (1996), «Mood, reasoning, and central executive processes», Journal of Experimental Psychology: Learning, Memory, and Cognition, 22, 477-493.

OAKSFORD, M.; CHATER, N., y LARKIN, J. (2000), «Probabilities and polarity biases in conditional inference», Journal of Experimental Psychology: Learning, Memory and Cognition, 26, 883-899.

OATLEY, K. (1976), «Inference, navigation and cognitive maps». En Jhonson-Laird, P. N., y Wason, P. C. (eds.), Thinking: Readings in Cognitive Science. Cambridge: Cambridge- University Press, pp. 537-547.

OATLEY, K. (1978), Perceptions and Representations: the Theoretical Bases of Brain Research and Psychology. Londres: Methuen.

OATLEY, K. (1996), «Emotions, rationality, and informal reasoning». En Oakhill, J., y Garnham, A. (eds.), Mental Models in Cognitive Science. Hove,

Sussex: Psychology Press, pp. 175-196.

OATLEY, K. (1998), A Natural History. Toronto, Ontario: Viking Penguin.

OATLEY, K., y JOHNSON-LAIRD, P. N. (1987), Towards a cognitive theory of emotion. Cognitionand Emotion, 1, 29-50.

OATLEY, K., y JOHNSON-LAIRD, P. N. (1996), «The communicative theory of emotions: empirical tests, mental models, and implications for social interaction». En Martin, L. L., y Tesser, A. (eds.), Striving and Feeling: Interactions Among Goals, Affect, and Self-regulation. Mahwah, NJ: Erlbaum, pp. 363-393.

OBERAUER, K., y WILHELM, O. (2000), «Effects of directionality in deductive reasoning: I. The comprehension of single relational premises», Journal of Experimental Psychology: Learning, Memory, and Cognition, 26, 1702-1712.

OBERAUER, K., y WILHELM, W. (2003), «The meaning(s), of conditionals: conditional probability, mental models, and personal utilities», Journal of Experimental Psychology: Learning, Memory & Cognition, 29, 680-693.

OBERAUER, K.; HÖRNIG, R.; WEIDENFELD, A., y WILHELM, O. (2005), «Effects of directionality in deductive reasoning, II: Premise integration and conclusion evaluation», Quarterly Journal of Experimental Psychology, 58A, 1225-1247.

OCHSNER, K. N., y GROSS, J. J. (2005), «The cognitive control of emotion», Trends in Cognitive Sciences, 9, 242-249.

O'CONNOR, K., y AARDEMA, F. (2003), «Fusion or confusion in obsessive-compulsive disorder», Psychological Reports, 93, 227-232.

OHLSSON, S. (1984a), «Restructuring revisited I: Summary and critique of Gestalt theory of problem solving», Scandavian Journal of Psychology, 25, 65-76.

OHLSSON, S. (1984b), «Restructuring revisited II: an information processing theory of restructuring and insight», Scandavian Journal of Psychology, 25, 117-129.

O’KEEFE, J., y NADEL, L. (1978), The Hippocampus as a Cognitive Map. Oxford: Clarendon Press.

O’NEILL, J. J. (1944), Prodigal Genius: the Life of Nikola Tesla. Nueva York: Ives Washburn.

ORLEANS, S., y ORLEANS, J. (1983), The Great Big Book of Pencil Puzzles. Nueva York: Pedigree.

ORMEROD, T. C.; MACGREGOR, J. N., y CHRONICLE, E. P. (2002), «Dynamics and constraints in insight problem solving», Journal of Experimental Psychology: Learning, Memory, and Cognition, 28, 791-799.

ORMEROD, T. C.; MANKTELOW, K. I., y JONES, G. V. (1993), «Reasoning with three types of conditional: biases and mental models», Quarterly Journal of Experimental Psychology, 46A, 653-678.

ORMEROD, T. C., y RICHARDSON, J. (2003), «On the generation and evaluation of inferences from single premises», Memory & Cognition, 31, 467-478.

ORMEROD, J. E. (1979), «Cognitive processes in the solution of three-term series problems», American Journal of Psychology, 92, 235-255.

ORTONY, A., y TURNER, T. J. (1990), «What’s basic about basic emotions?», Psychological Review, 74, 431-461.

ORLONY, A.; CLORE, G. I., y COLLINS, A. (1988), The Cognitive Structure of Emotions New York: Cambridge University Press.

O’SHEA, S. (2000), The Perfect Heresy: the Revolutionary Life and Death of the Medieval Cathars. Nueva York: Walker.

OSHERSON, D. N. (1974-76), Logical Abilities in Children, Vols 1-4. Hillsdale, NJ: Erlbaum.

OSHERSON, D. N.; SMITH, E. E., y SHAFIR, E. (1986), «Some origins of belief», Cognition, 24, 197-224.

OSHERSON, D.; PERANI, D.; CAPPA, S.; SCHNUR, T.; GRASSI, F., Y

FAZIO, F. (1998), «Distinct brain loci in deductive versus probabilistic reasoning», Neuropsychologia, 36, 369-376.

OSOFSKY, M. J.; BANDURA, A., y ZIMBARDO, P. G. (2005), «The role of moral disengagement in the execution process», Law and Human Behavior, 29, 371-393.

OSWALD, N. C.; WALLER, R. E., y DRINKWATER, J. (1970), «Relationship between breathlessness and anxiety in asthma and bronchitis: a comparative study», British Medical Journal, 2, 14-17.

OTERO, J., y KINTSCH, W. (1992), «Failures to detect contradictions in a text: what readers believe versus what they read», Psychological Science, 3, 229-235.

OVER, D. E., y EVANS, J. St. B. T. (1997), «Two cheers for deductive competence», Current Psychology of Cognition, 16, 225-278.

PACCANARO, A., y HINTON, G. E. (2000), «Learning distributed representations of concepts from relational data using linear relational embedding», IEEE Transactions on Knowledge and Data Engineering, 13, 232-245.

PAINE, T. (1989), Political Writings. Kucklik, B. (ed.). Cambridge: Cambridge University Press.

PAIVIO, A. (1971), Imagery and Verbal Processes. Nueva York: Holt, Rinehart & Winston.

PARIS, S. G. (1973), «Comprehension of language connectives and propositional logical relationships», Journal of Experimental Child Psychology, 16, 278-291.

PARSONS, L. M., y OSHERSON, D. (2001), «New evidence for distinct right and left brain systems for deductive versus probabilistic reasoning», Cerebral Cortex, 11, 954-965.

PARTEE, B. H. (1995), «Lexical semantics and compositionality». En Gleitman, L. R., y Liberman, M. (eds.), Language: an Invitation to Cognitive Science, Vol. 1. Cambridge, MA: MIT Press, pp. 311-360.

PASCAL, B. (1966), Pensées (trad. Krailsheimer, A. J.). Harmonsworth, Middx, UK: Penguin. (Primera versión publicada en 1670.)

PEARL, J. (1988), Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Segunda edición revisada. San Francisco: Morgan Kaufmann.

PEARL, J. (2000), Causality: Models, Reasoning, and Inference. Cambridge: Cambridge University Press.

PEARSON, H. (1984), The Smith of Smiths: Being the Life, Wit, and Humour of Sydney Smith. Londres: Hogarth. (Publicado originalmente en 1934.)

PEIRCE, C. S. (1931-1958), Hartshorne, C.; Weiss, P., y Burks, A. (eds.), Collected Papers of Charles Sanders Peirce, 8 vols. Cambridge, MA: Harvard University Press.

PEIRCE, C. S. (1955), Buchler, J. (ed.), Philosophical Writings of Peirce. Nueva York: Dover.

PELISSIER, M.-C., y O'CONNOR, K. P. (2002), «Deductive and inductive reasoning in obsessive- compulsive disorder», British Journal of Clinical Psychology, 41, 15-27.

PENG, K., y NISBETT, R. E. (1999), «Culture, dialectics, and reasoning about contradiction», American Psychologist, 54, 741-754.

PERHAM, N., y OAKSFORD, M. (2005), «Deontic reasoning with emotional content: evolutionary psychology or decision theory?», Cognitive Science, 29, 681-718.

PERKINS, D. N. (1981), The Mind's Best Work. Cambridge, MA: Harvard University Press.

PERKINS, D. (2000), The Eureka Efffect: the Art and Logic of Breakthrough Thinking. Nueva York: Norton.

PERKINS, D. N.; FARADY, M., y BUSHEY, B. (1990), «Everyday reasoning and the roots of intelligence». En Voss, J.; Perkins, D. N., y Segal, J. (eds.), Informal Reasoning and Education. Hillsdale, NJ: Erlbaum, pp. 83-105.

PERNER, J. (1991), Understanding the Representational Mind. Cambridge, MA: MIT Press.

PERNER, J.; LEEKHAM, S., y WIMMER, H. (1987), «Three year olds' difficulty with false belief: the case for conceptual deficit», British Journal of Developmental Psychology, 5, 125-137.

PERNER, J.; SPRUNG, M., y STEINKOGLER, B. (2004), «Counterfactual conditionals and false belief: a developmental dissociation», Cognitive Development, 19, 179-201.

PETERSON, D., y RIGGS, K. J. (1999), «Adaptive modelling and mindreading», Mind & Language, 14, 80-112.

PIAGET, J. (1928), Judgment and Reasoning in the Child. Londres: Kegan Paul.

PIAGET, J. (1965a), Logic and Psychology. [Ed. cast.: Estudios sobre lógica y psicología, trad. de A. Deaño y J. Delval, Madrid, Alianza, 1982.]

PIAGET, J. (1965b), The Moral Judgment of the Child. Nueva York: Free Press.

PIAGET, J. (1987), Possibility and Necessity, Vols. 1 y 2. Minneapolis, MN: University of Minnesota Press.

PIAGET, J., e INHELDER, B. (1969), The Psychology of the Child. Nueva York: Basic Books.

PICA, P.; LEMER, C.; IZARD, V., y DEHAENE, S. (2004), «Exact and approximate arithmetic in an Amazonian indigene group. Science, 496, 499-503.

PINKER, S. (1997), How the Mind Works. Nueva York: W. W. Norton.

PINKER, S. (2002), The Blank Slate: the Modern Denial of Human Nature. Nueva York: Viking.

PLATÓN (1961), Hamilton, E., y Cairns, H. (eds.), Complete Works: the Collected Dialogues of Plato. Princeton, NJ: Princeton University Press.

PLATT, R. D., y GRIGGS, R. A. (1993), «Facilitation in the abstract selection task: the effects of attentional and instructional factors», Quarterly Journal of

Experimental Psychology, 46A, 591-613.

POE, E. A. (1982), The Complete Tales and Poems of Edgar Allan Poe. Londres: Penguin.

POLITZER, G. (1986), Laws of language use and of formal logic. Journal of Psycholinguistic Research, 15, 47-92.

POLITZER, G., y BOURMAUD, G. (2002), «Deductive reasoning from uncertain conditionals», British Journal of Psychology, 93, 345-381.

POLITZER, G., y CARLES, L. (2001), «Belief revision and uncertain reasoning», Thinking & Reasoning, 7 217-234.

POLK, T. A., y NEWELL, A. (1995), «Deduction as verbal reasoning», Psychological Review, 102, 533-566.

POLLOCK, J. (1989), How to Build a Person: a Prolegomenon. Cambridge, MA: MIT Bradford Books.

POLYA, G. (1973), How To Solve It: a New Aspect of Mathematical Methods, 2ª ed. Princeton, NJ: Princeton University Press. (Publicado originalmente en 1945.)

POPPER, K. (1959), The Logic of Scientific Discovery. Londres: Hutchinson.

PORTER, R. (1997), The Greatest Benefit to Mankind: a Medical History of Humanity. Nueva York: W. W. Norton.

POTTS, G. R., y SCHOLZ, K. W. (1975), «The internal representation of a three-term series problem», Journal of Verbal Learning and Verbal Behavior, 14, 439-452.

POUNDSTONE, W. (1988), Labyrinths of Reason: Paradox, Puzzles, and the Frailty of Knowledge. Nueva York: Anchor Books, Uoubleday.

POWER, M., y DALGLEISH, T. (1997), Cognition and Emotion: From Order lo Disorder. Hove, East Sussex: Psychology Press.

POWER, M., y WYKES, T. (1996), «The mental health of mental models and

the mental models of mental health». En Oakhill, J., y Garnham, A. (eds.), Mental Models in Cognitive Science. Hove, Sussex: Psychology Press, pp. 197-222.

POWER, R. (1984), Inverted Deductions with Numerical Quantifiers. Report 104. Padua: Universidad de Padua, Instituto de Psicología.

POWERS, A. L. (ed.), (1997), Italy in Mind: an Anthology. Nueva York: Random House.

PRAWITZ, D. (1965), Natural Deduction: a Proof-theoretical Study. Uppsala, Suecia: Almqvist y Wiksell.

PREMACK, D. (1990), «Words: What are they, and do animals have them?», Cognition, 37, 197-212.

PYLYSHYN, Z. (1973), «What the mind's eye tells the mind's brain: a critique of mental imagery», Psychological Bulletin, 80, 1-24.

PYLYSHYN, Z. (1981), «The imagery debate: analogue media versus tacit knowledge», Psychological Review, 88, 16-45.

PYLYSHYN, Z. (2003), «Return of the mental image: Are there really pictures in the head?», Trends in Cognitive Sciences, 2003, 7, 113-118.

QUELHAS, C., y BYRNE, R. M. J. (2003), «Reasoning with deontic and counterfactual conditionals», Thinking & Reasoning, 9, 43-66.

QUINE, W. V. O. (1952), Methods of Logic. Londres: Routledge & Kegan Paul.

QUINE, W. V. O. (1955), «A way to simplify truth functions», American Mathematical Monthly, 59, 521-531.

QUINE, W. V. O. (1974), Methods of Logic, 3ª ed. Londres: Routledge.

QUINN, S., y MARKOVITS, H. (2002), «Conditional reasoning with causal premises: evidence for a retrieval model», Thinking & Reasoning, 8, 179-191.

QUINTON, G., y FELLOWS, B. J. (1975), «"Percepual" strategies in the solving of three-term series problems», British journal of Psychology, 66, 69-78.

RABBITT, P. M. A., y ANDERSON, M. (2006), «The lacunae of loss? Aging and the differentiation of cognitive abilities». En Bialystok, E., y Craik, F. I. M. (eds.), Lifespan Cognition: Mechanisms of Change. Oxford: Oxford University Press. En prensa.

RADER, A. W., y SLOUTSKY, V. M. (2001), «Conjunction bias in memory representations of logical connectives», Memory & Cognition, 29, 838-849.

RADER, A. W., y SLOUTSKY, V. M. (2002), «Processing of logically valid and logically invalid conditional inferences in discourse comprehension», Journal of Experimental Psychology: Learning, Memory and Cognition, 28, 59-68.

RAGNI, M. (2003), «An arrangement calculus, its complexity and algorithmic properties», Advances in Artificial Intelligence: Lecture Notes in Computer Science, 2821, 580-590.

RAMSEY, F. P. (1990), Mellor, D. H. (ed.), Foundations: Essays in Philosophy, Logic, Mathematics and Economics. Londres: Humanities Press, pp. 145-163. (Publicado originalmente en 1929.)

REASON, J. T. (1988), «Stress and cognitive failure». En Fisher, S., y Reason, J. T. (eds.), Handbook of Life Stress, Cognition and Health. Chichester: Wiley, pp. 405-421.

REASON, J. T. (1990), Human Error. Cambridge: Cambridge University Press.

REED, G. (1977), «Obsessional cognition: performance on two numerical tasks», British Journal of Psychiatry, 130, 184-185.

REHDER, B., y HASTIE, R. (1996), «The moderating influence and variability on belief revision», Psychonomic Bulletin ang Review, 3, 499-503.

REICHENBACH, H. (1956), The Direction of Time. Berkeley: University of California Press.

REID, T. (1846), The Works of Thomas Reid. Edimburgo: Maclachlan y Stewart.

REINHART, T. (1986), «On the interpretation of “donkey”-sentences». En Traugott, B. C.; ter Meulen, A.; Reilly, J. S., y Ferguson, C. A. (eds.), On

Conditionals. Cambridge: Cambridge University Press, pp. 103-122.

REITER, R. (1980), «A logic for default reasoning», Artificial Intelligence, 13, 81-132.

REITMAN, W. R. (1965), Cognition and Thought. Nueva York: Wiley.

RESCHER, N. (1973), The Coherence Theory of Truth. Oxford: Oxford University Press.

REVLIN, R.; CATE, C. L., y ROUSS, T. S. (2001), «Reasoning counterfactually: combining and rending», Memory & Cognition, 29, 1196-1208.

REVLIS, R. (1974), «Prevarication: Reasoning from false assumptions», Memory & Cognition, 2A, 87-95.

REVLIS, R. (1975), «Two models of syllogistic reasoning: feature selection and conversion», Journal of Verbal Learning and Verbal Behavior, 14, 180-195.

REVLIS, R., y HAYES, J. R. (1972), «The primacy of generalities in hypothetical reasoning», Cognitive Psychology, 3, 268-290.

REVLIS, R.; LIPKIN, S. G., y HAYES, J. R. (1971), «The importance of universal quantifiers in a hypothetical reasoning task», Journal of Verbal Learning and Verbal Behavior, 10, 86-91.

RHÉAUME, J.; FREESTON, M. H.; LADOUCEUR, R.; BOUCHARD, C.; GALLANT, L.; TALBOT, F., y VALLIÉRES, A. (2000), «Functional and dysfunctional perfectionists: are they different on compulsive-like behaviors?», Behaviour Research and Therapy, 38, 119-128.

RIBE, N., y STEINLE, F. (2005), «Exploratory experimentation: Goethe, Land, and color theory», www.physicstoday.org/vol-55/iss-7/p43.html.

RICHARDSON, J. T. E. (1987), «The role of mental imagery in models of transitive inference», British Journal of Psychology, 78, 189-203.

RICHARDSON, J., y ORMEROD, T. C. (1997), «Rephrasing between disjunctives and conditionals: mental models and the effects of thematic

content», Quarterly Journal of Experimental Psychology, 50A, 358-385.

RIESBECK, C. K., y SCHANK, R. C. (1989), Inside Case-Based Reasoning. Hillsdale, NJ: Erlbaum.

RIGGS, K. J.; PETERSON, D. M.; ROBINSON, E. J., y MITCHELL, P. (1998), «Are errors in false belief tasks symptomatic of a broader difficulty with counterfactuality?», Cognitive Development, 13, 73-90.

RILEY, C. A. (1976), «The representation of comparative relations and the transitive inference task», Journal of Experimental Child Psychology, 22, 1-22.

RILEY, C. A., y TRABASSO, T. (1974), «Comparatives, logical structures, and encoding in a transitive inference task», Journal of Experimental Child Psychology, 17, 187-203.

RIPS, L. J. (1983), «Cognitive processes in prepositional reasoning», Psychological Review, 90, 38-71.

RIPS, L. J. (1989), «The psychology of knights and knaves», Cognition, 31, 85-116.

RIPS, L. J. (1990), «Paralogical reasoning: Evans, Johnson-Laird, y Byrne on liar and truth-teller puzzles», Cognition, 36, 291-314.

RIPS, L. J. (1994), The Psychology of Proof. Cambridge, MA: MIT Press.

RIPS, L. J. (1998), «Reasoning and conversation», Psychological Review, 105, 411-441.

RIPS, L. J. (2002), «Reasoning». En Medin, D. (ed.), Stevens' Handbook of Experimental Psychology, Vol. 2: Memory and Cognitive Processes, 3ª ed. Nueva York: John Wiley, pp. 317-362.

RIPS, L. J., y COLLINS, A. (1993), «Categories and resemblance», Journal of Experimental Psychology: General, 122, 468-486.

RIPS, L. J.; SHOBEN, E. J., y SMITH, E. E. (1973), «Semantic distance and the verification of semantic relations», Journal of Verbal Learning and Verbal Behavior, 12, 1-20.

RIPS, L. J.; BREM, S. K., y BAILENSON, J. N. (1999), «Reasoning dialogues», Current Directions in Psychological Science, 8, 172-177.

ROBERTS, M. J. (2000), «Strategies in relational inference», Thinking & Reasoning, 6, 1-26.

ROBERTS, M. J. (2006), «Falsification and mental models: it depends on the task». En Schaeken, W.; Vandierendonck, A.; Schroyens, W., y d'Ydewalle, G. (eds.), The Mental Models Theory of Reasoning: Refinement and Extensions. Mahwah, NJ: Erlbaum. En prensa.

ROBERTS, M. J., y SYKES, E. D. A. (2005), «Categorical reasoning from multiple diagrams», Quarterly Journal of Experimental Psychology, 58A, 333-376.

ROBINSON, L. A.; BERMAN, J. S., y NEIMEYER, R. A. (1990), «Psychotherapy for the treatment of depression: a comprehensive review of controlled outcome research», Psychological Bulletin, 108, 30-49.

ROESE, N. J. (1997), «Counterfactual thinking», Psychological Bulletin, 121, 133-148.

ROGERS, L. J. (1997), Minds of Their Own. St. Leonards, UK: Allen & Unwin.

ROSCH, E. (1973), «Natural categories», Cognitive Psychology, 4, 328-35.

ROSENBLATT, F. (1958), «The perceptron: a probabilistic model for information storage and organization in the brain», Psychological Review, 65, 386-408.

ROSENBLATT, F. (1961), Principles ofNeurodynamics. Washington, DC: Spartan Press.

ROSS, L. (1977), «The intuitive psychologist and his shortcomings: distortions in the attribution process». En Berkowitz, L. (ed.), Advances in Experimental Social Psychology, Vol. 10. Nueva York: Academic Press, pp. 174-221.

ROSS, L., y LEPPER, M. R. (1980), «The perseverance of beliefs: empirical and normative considerations». En Shweder, R. A. (ed.), Fallible Judgement in Behavioral Research: New Directions for Methodology of Social and Behavioral

Science, Vol. 4. San Francisco, CA: Jossey-Bass, pp. 17-36.

ROYZMAN, E. B., y BARON, J. (2002), «The preference for indirect harm», Social Justice Research, 15, 165-184.

ROZENBLIT, L. R., y KEIL, F. C. (2002), «The misunderstood limits of folk science: an illusion of explanatory depth», Cognitive Science, 26, 521-562.

ROZIN, P. (1996), «Towards a psychology of food and eating: from motivation to module to model to marker, morality, meaning, and metaphor», Current Directions in Psychological Science, 5, 18-24.

RUSSELL, B. A. W. (1905), «On denoting», Mind, 14, 479-493.

RUSSELL, B. A. W. (1912-13), «On the notion of cause», Proceedings of the Aristotelian Society, 13, 1-26.

RUSSELL, B. A. W. (1927), An Outline of Philosophy. Londres: Allen and Unwin.

RUSSELL, B. A. W. (1946), The History of Western Philosophy. Nueva York: Simon & Schuster.

RUSSELL, B. A. W. (1957), «Knowledge by acquaintance and knowledge by description». En su Mysticism and Logic. Garden City, NY: Uoubleday. (Publicado originalmente en 1917.)

RUSSELL, B. A. W. (1967), The Autobiography of Bertrand Russell, 3 vols. Boston: Little, Brown.

RUSSELL, J. (1987), «Rule-following, mental models, and the developmental view». En Chapman, M., y Dixon, R. A. (eds.), Meaning and the Growth of Understanding: Wittgenstein’s Significance for Developmental Psychology. Nueva York: Springer, pp. 23-48.

RUSSELL, J. A. (1991), «In defense of a prototype approach to emotion concepts», Journal of Personality and Social Psychology, 60, 37-47.

RUSSELL, S. (1989), The Use of Knowledge in Analogy and Induction. Londres: Pitman.

RUSSO, J. E.; JOHNSON, E. J., y STEPHENS, D. L. (1989), «The validity ol verbal protocols», Memory & Cognition, 17, 759-769.

SAINSBURY, M. (2001), Logical Forms: an Introduction to Philosophical Logic, 2ª ed. Oxford: Blackwell.

SALKOVSKIS, P. M. (1996), «The cognitive approach to anxiety: threat beliefs, safety-seeking behavior, and the special case of health anxiety and obsession». En Salkovskis, P. M. (ed.), Frontiers of Cognitive Therapy. Nueva York: Guilford Press, pp. 48-74.

SALMON, W. C. (1980), «Probabilistic causality», Pacific Philosophical Quarterly, 61, 50-74.

SALSBURG, D. (2001), The Lady Tasting Tea: How Statistics Revolutionized Science in the Twentieth Century. Nueva York: W. H. Freeman.

SAMARAPUNGAVAN, A.; VOSNIADOU, S., y BREWER, W. (1996), «Mental models of the earth, sun and moon: Indian children's cosmologies», Cognitive Development, 11, 491-521.

SANTAMARIA, C., y JOHNSON-LAIRD, P. N. (2000), «An antidote to illusory inferences», Thinking & Reasoning, 6, 313-333.

SARTRE, J.-P. (1938), Esquisse d'une théorie des émotions. [Ed. cast.: Bosquejo de una teoría de las emociones, trad. de M. Acheroff, Madrid, Alianza, 2005.]

SAVAGE, L. J. (1954), The Foundations of Statistics. Nueva York: Wiley.

SCHACHTER, S., y SINGER, J. E. (1962), «Cognitive, social, and psychological determinants of emotional state», Psychological Review, 69, 379-399.

SCHAEKEN, W. S.; JOHNSON-LAIRD, P. N., y D'YDEWALLE, G. (1996a), «Mental models and temporal reasoning», Cognition, 60, 205-234.

SCHAEKEN, W. S.; JOHNSON-LAIRD, P. N., y D'YDEWALLE, G. (1996b), «Tense, aspect, and temporal reasoning», Thinking and Reasoning, 2, 309-327.

SCHAEKEN, W. S.; GIROTTO, V., y JOHNSON-LAIRD, P. N. (1998), «The

effect of an irrelevant premise on temporal and spatial reasoning», Kognitionswisschenschaft, 7, 27-32.

SCHAEKEN, W. S.; DE VOOGHT, G.; VANDIERENDONCK, A., y D'YDEWALLE, G. (2000), Deductive Reasoning and Strategies. Mahwah, NJ: Erlbaum.

SCHAEKEN, W. S.; VAN DER HENST, J.-B., y SCHROYENS, W. (en prensa), «The mental models theory of relational reasoning: premise relevance, conclusion phrasing and cognitive economy». En Schaeken, W.; Vandierendonck, A.; Schroyens, W., y d'Ydewalle, G. (eds.), The Mental Models Theory of Reasoning: Extensions and Refinements. Mahwah, NJ: Erlbaum.

SCHANK, R. C., y ABELSON, R. P. (1977), Scripts, Plans, Goals and Understanding. Hillsdale, NJ: Erlbaum.

SCHKADE, D. A., y KAHNEMAN, D. (1998), «Does living in California make people happy?», Psychological Science, 9, 340-346.

SCHLOTTMANN, A., y ANDERSON, N. H. (1995), «Belief revision in children: Serial judgment in social cognition and decision-making domains», Journal of Experimental Psychology: Learning, Memory, and Cognition, 21, 1349-1364.

SCHOOLER, J.; OHLSSON, S., y BROOKS, K. (1993), «Thoughts beyond words: when language overshadows insight», Journal of Experimental Psychology: General, 122, 166-183.

SCHROYENS, W. (1997), «Meta-propositional reasoning about the truth or falsity of propositions», Psychologica Belgica, 37, 219-247.

SCHROYENS, W., y SCHAEKEN, W. (2003), «A critique of Oaksford, Chater, and Larkin's (2000), conditional probability model of conditional reasoning», Journal of Experimental Psychology: Learning, Memory, and Cognition, 29, 1, 140-149.

SCHROYENS, W., y SCHAEKEN, W. (2004), «Guilt by association: on iffy propositions and the proper treatment of mental-models theory», Current Psychology Letters, 12, 1.

SCHROYENS, W.; SCHAEKEN, W., y D'YDEWALLE, G. (1996), «Meta-propositional reasoning with knight-knave problems: the importance of being hypothesized», Psychologica Belgica, 36, 145-169.

SCHROYENS, W.; SCHAEKEN, W., y D'YDEWALLE, G. (1999), «Error and bias in meta-propositional reasoning: a case of the mental model theory», Thinking & Reasoning, 5, 29-65.

SCHROYENS, W.; SCHAEKEN, W., y D'YDEWALLE, G. (2001a), «The processing of negations in conditional reasoning: a meta-analytical case study in mental models and/or mental logic theory», Thinking & Reasoning, 7, 121-172.

SCHROYENS, W.; SCHAEKEN, W., y D'YDEWALLE, G. (2001b), «A meta-analytic review of conditional reasoning by model and/or rule: mental models theory revised». Artículo sin publicar.

SCHROYENS, W.; SCHAEKEN, W., y HANDLEY, S. (2003), «In search of counter examples: deductive rationality in human reasoning», Quarterly Journal of Experimental Psychology, 56A (7), 1129-1145.

SCHUSTACK, M. W. (1988), «Thinking about causality». En Sternberg, R. J., y Smith, E. E. (eds.), The Psychology of Thinking. Cambridge: Cambridge University Press, pp. 92-115.

SCODITTI, G. M. G. (1984), «The use of "metaphors" in Kitawa culture, Northern Massim», Oceania, 55, 50-70.

SCOVILLE, W. B., y MILNER, B. (1957), «Loss of recent memory after bilateral hippocampal lesions», Journal of Neurology, Neurosurgery and Psychiatry, 20, 11-12.

SCRIBNER, S. (1977), «Modes of thinking and ways of speaking: culture and logic reconsidered». En Johnson-Laird, P. N., y Wason, P. C. (eds.), Thinking: Readings in Cognitive Science. Cambridge: Cambridge University Press, pp. 483-500.

SCRIVEN, M. (1976), Reasoning. Nueva York: McGraw-Hill.

SEARLE, J. R. (1980), «Minds, brains, and programs», Behavioral and Brain Sciences, 3, 417-424.

SEBEOK, T. A., y UMIKER-SEBEOK, J. (1983), «"You know my method": a juxtaposition of Charles S. Peirce and Sherlock Holmes». En Eco, U., y Sebeok, T. A. The Sign of Three: Dupin, Holmes, Peirce. Bloomington, IN: Indiana University Press, pp. 11-54.

SELIGMAN, M. E. P. (1995), «The effectiveness of psychotherapy: the Consumer Reports Study», American Psychologist, 50, 965-974.

SELLS, S. B. (1936), «The atmosphere effect: an experimental study of reasoning», Archives of Psychology, 29, 3-72.

SHAFER, G. (1996), The Art of Causal Conjecture. Cambridge: MIT Press.

SHAFIR, E., y TVERSKY, A. (1992), «Thinking through uncertainty: nonconsequential reasoning and choice», Cognitive Psychology, 24, 449-474.

SHAFIR, S. (1994), «Intransitivity of preferences in honey bees—support for comparative evaluation of foraging options», Animal Behaviour, 48, 55-67.

SHANKS, D. R. (2004), «Judging covariation and causation». En Koehler, D. J., y Harvey, N. (eds.), Handbook of Judgment and Decision Making. Oxford: Blackwell, pp. 220-239.

SHANNON, C. E., y WEAVER, W. (1949), The Mathematical Theory of Communication. Urbana, IL: University of Illinois Press.

SHASTRI, L. (1988), «A connectionist approach to knowledge representation and limited inference», Cognitive Science, 12, 331-392.

SHASTRI, L., y AJJANAGADDE, V. (1993), «From simple associations to systematic reasoning: a connectionist representation of rules, variables and dynamic bindings using temporal synchrony», Behavioral and Brain Sciences, 16, 417-494.

SHAVER, P. R.; PIERSON, L., y LANG, S. (1975), «Converging evidence for the functional significance of imagery in problem solving», Cognition, 3, 359-375.

SHAW, B. (1965), The Complet Plays of Bernard Shaw. Londres: Paul Hamlym.

SHAW, V. K., y JOHNSON-LAIRD, P. N. (1998), «Dispelling the “atmosphere” effect in reasoning», Analise Psicologia. Special issue on Cognition and Context, Quelhas, A. C., y Pereira, K. (eds.), 169-199.

SHEEN, Mr Justice (1987), «Report of Court No. 8074; mv Herald of Free Enterprise». Londres: Her Majesty’s Stationary Office (HMSO).

SHEPARD, R. N., y METZLER, J. (1971), «Mental rotation of three-dimensional objects», Science, 171, 701-703.

SHEPARD, R. N.; HOVLAND, C. I., y JENKINS, H. M. (1961), «Learning and memorization of classifications», Psychological Monographs: General and Applied, 75, 1-42.

SHIH, M.; PITTINSKY, T. L., y AMBADY, N. (1999), «Stereotype threat and women’s math performance», Psychological Science, 10, 80-83.

SHIMOJO, S., y ICHIKAWA, S. (1989), «Intuitive reasoning about probability: theoretical and experimental analyses of the “problem of three prisoners”», Cognition, 32, 1-24.

SHINN, S.-J. (1994), «Peirce and the logical status of diagrams», History and Philosophy of Logic, 15, 45-68.

SHORTLIFFE, E. H. (1976), Computer-Based Medical Consultation: MYCIN. Nueva York: Elsevier.

SIEG, W. (1999), «Church-Turing thesis». En Wilson, R. A., y Keil, E. C. (eds.), The MIT Encyclopedia of the Cognitive Sciences. Cambridge, MA: MIT, pp. 116-118.

SIFNEOS, P. E. (1996), «Alexithymia: past and present», American Journal of Psychiatry, 153 (7 Supl.), 137-142.

SIMON, H. A. (1982), Models of Bounded Rationality, Vols. 1 y 2. Cambridge, MA: MIT Press.

SIMON, H. A. (1983), «Search and reasoning in problem solving», Artificial Intelligence, 21, 7-29.

SIMON, H. A. (1991a), Models of My Life. Nueva York: Basic Books.

SIMON, H. A. (1991b), «Cognitive architectures and rational analysis: comment». En VanLehn, K. (ed.), Architectures for Intelligence: the 22nd Carnegie Mellon Symposium on Cognition. Hillsdale, NJ: Erlbaum.

SIMONTON, D. K. (1995), «Foresight in insight? A Darwinian answer». En Sternberg, R. J., y Davidson, J. E. (eds.), The Nature of Insight. Cambridge, MA: MIT Press, pp. 465-494.

SKINNER, B. F. (1953), Science and Human Behavior. Nueva York: Macmillan.

SLOMAN, S. A. (1996), «The empirical case for two systems of reasoning», Psychological Bulletin, 119, 3-22.

SLOMAN, S. A. (2005), Causal Models: How People Think about the World and Its Alternatives. Nueva York: Oxford University Press.

SLOMAN, S. A., y LAGNADO, D. A. (2005), «Do we “do”?», Cognitive Science, 29, 5-39.

SLOMAN, S. A.; OVER, D.; SLOVAK, L., y STIBEL, J. (2003), «Frequency illusions and other fallacies», Organizational Behavior and Human Decision Processes, 91, 296-309.

SLOUTSKY, V. M., y GOLDVARG, Y. (2004), «Mental representation of logical connectives», Quarterly Journal of Experimental Psychology, 57A, 636-665.

SMEETS, G., y DE JONG, P. J. (2005), «Belief bias and symptoms of psychopathology in a non-clinical sample», Cognitive Therapy and Research, 29, 377-386.

SMITH, E. E.; LANGSTON, C., y NISBETT, R. E. (1992), «The case for rules in reasoning», Cognitive Science, 16, 1-40.

SMITH, M. L.; GLASS, G. V., y MILLER, T. I. (1980), The Benefits of Psychotherapy. Baltimore, MD: Johns Hopkins University Press.

SMITH, S. M., y BLANKENSHIP, S. E. (1991), «Incubation and the persistence

of fixation in problem solving», American Journal of Psychology, 104, 61-87.

SMOLCESKY, P. (1990), «Tensor product variable binding and the representation of symbolic structures in connectionist systems», Artificial Intelligence, 46, 159-216.

SMULLYAN, R. M. (1978), What is the Name of This Rook? The Riddle of Dracula and Other Logical Puzzles. Englewood Cliffs, NJ: Prentice-Hall.

SNOW, J. (1849a), On the Mode of Communication of Cholera. Londres: Churchill.

SNOW, J. (1849b), «On the pathology and mode of communication of cholera», London Medical Gazette, 44, 745-752, 923-929.

SNOW, J. (1855), On the Mode of Communication of Cholera, 2ª ed. Londres: Churchill.

SOWA, J. F. (1984), Conceptual Structures: Information Processing in Mind and Machine. Reading, MA: Addison-Wesley.

SOWA, J. F., y MAJUMDAR, A. K. (2003), «Analogical reasoning». En Aldo, A.; Lex, W., y Ganter, B. (eds.), Conceptual Structures for Knowledge Creation and Communication, LNAl 2746. Berlin: Springer-Verlag, pp. 16-36.

SPELKE, E. S. (1990), «Principles of object permanence», Cognitive Science, 14, 29-56.

SPELKE, E. S. (1991), «Physical knowledge in infancy: reflections on Piaget's theory». En Carey, S., y Gelman, R. (eds.), The Epigenesis of Mind: Essays on Biology and Cognition. Hillsdale, NJ: Erlbaum, pp. 133-169.

SPENDER, S. (1952), «The making of a poem». En Ghiselin, B. (ed.), The Creative Process. New York: Mentor, pp. 112-125.

SPERBER, D. (1996), Explaining Culture. Oxford: Blackwell.

SPERBER, D., y WILSON, D. (1986), Relevance: Communication and Cognition. Oxford: Basil Blackwell.

SPERBER, D.; CARA, F., y GIROTTO, V. (1995), «Relevance theory explains the selection task», Cognition, 52, 3-39.

SPIRTES, P.; GLYMOUR, C., y SCHEMES, R. (1993), Causation, Prediction, and Search. Nueva York: Springer-Verlag.

STALNAKER, R. C. (1968), «A theory of conditionals». En Rescher, N. (ed.), Studies in Logical Theory (American Philosophical Quarterly Monograph No. 2). Oxford: Blackwell, pp. 98-122.

STANFILL, C., y WALTZ, D. (1986), «Toward memory-based reasoning», Communications of the Association for Computing Machinery, 29, 1213-1228.

STANOVICH, K. E. (1999), Who is Rational? Studies of Individual Differences in Reasoning. Mahwah, NJ: Erlbaum.

STANOVICH, K. E., y WEST, R. F. (1998), «Cognitive ability and variation in selection task performance», Thinking & Reasoning, 4, 193-230.

STEELE, C., y ARONSON, J. (1995), «Stereotype threat and intellectual test performance of African Americans», Journal of Personality and Social Psychology, 69, 797-811.

STEIN, N. L.; TRABASSO, T., y LIWAG, M. (1993), «The representation and organization of emotional experience: unfolding the emotion episode». En Lewis, M., y Haviland, J. M. (eds.), Handbook of Emotions, 1ª ed. Nueva York: Guilford Press, pp. 279-300.

STEINMEYER, J. (2003), Hiding the Elephant: How Magicians Invented the Impossible and Learned to Disappear. Nueva York: Carroll and Graf.

STENNING, K. (2002), Seeing Reason: Image and Language in Learning to Think. Oxford: Oxford University Press.

STENNING, K., y VAN LAMBALGEN, M. (2005), «Semantic interpretation as computation in nonmonotonic logic: the real meaning of the suppression task», Cognitive Science, 29, 919-960.

STENNING, K., y OBERLANDER, J. (1995), «A cognitive theory of graphical and linguistic reasoning: logic and implementation», Cognitive Science, 19, 97-

140.

STENNING, K., y YULE, P. (1997), «Image and language in human reasoning: a syllogistic illuatration», Cognitive Psychology, 34, 109-159.

STERN, J. (2003), Terror in the Name of God: Why Religious Militants Kill. Nueva York: Ecco.

STERNBERG, R. J. (1977), Intelligence, Information Processing, and Analogical Reasoning: the Componential Analysis of Human Abilities. Hillsdale, NJ: Erlbaum.

STERNBERG, R. J. (1980), «Representation and process in linear syllogistic reasoning», Journal of Experimental Psychology: General, 10, 119-159.

STERNBERG, R. J. (1983), «Components of human intelligence», Cognition, 15, 1-48.

STERNBERG, R. J. (ed.), (1984), Human Abilities: an Information-Processing Approach. San Francisco, CA: Freeman.

STERNBERG, R. J. (1997), Thinking Styles. Nueva York: Cambridge University Press.

STERNBERG, R. J., y WEIL, E. M. (1980), «An aptitude-strategy interaction in linear syllogistic reasoning», Journal of Educational Psychology, 72, 226-239.

STEVENSON, R. J, y OVER, D. E. (1995), «Deduction from uncertain premises», Quarterly Journal of Experimental Psychology, 48A, 613-643.

STEVENSON, R. J., y OVER, D. E. (2001), «Reasoning from uncertain premises: effects of expertise and conversational context», Thinking & Reasoning, 7, 367-390.

STICH, S. P. (1983), From Folk Psychology to Cognitive Science: the Case against Belief. Cambridge, MA: MIT Press.

STÖRRING, G. (1908), «Experimentelle Untersuchungen über einfache Schlussprozesse», Archiv fur die gesamte Psychologie, 11, 1-27.

STRAWSON, P. F. (1950), «On referring», Mind, 59, 320-344.

SUBER, P. (1990), The Paradox of Self-Amendment: a Study of Logic, Law, Omnipotence, and Chance. Berlin: Peter Lang.

SUNSTEIN, C. R. (2005), «Moral heuristics», Behavioral and Brain Sciences, 28, 531-573.

SUPPES, P. (1970), A Probabilistic Theory of Causality. Amsterdam: North-Holland.

SUPPES, P. (1984), Probabilistic Metaphysics. Oxford: Basil Blackwell.

SUTHERLAND, J. (1996), Is Heathcliff a Murderer: Great Puzzles in Nineteenth-Century Literature. Nueva York: Oxford University Press.

SUTHERLAND, N. S. (1976), Breakdown: a Personal Crisis and a Medical Dilemma. Londres: Weidenfeld & Nicolson.

SZASZ, T. (1974), The Myth of Mental Illness: Foundations of a Theory of Personal Conduct, edn revisada. Nueva York: Harper.

TABACHNECK, H., y SIMON, H. (1992), «Effect of mode of presentation on reasoning about economic markets». En Narayanan, N. H. (ed.), AAAI Spring Symposium on Reasoning with Diagrammatic Representations, 25-27 de marzo, Universidad de Stanford California, pp. 56-64.

TABOSSI, P.; BELL, V. A., y JOHNSON-LAIRD, P. N. (1998), «Mental models in deductive, modal, and probabilistic reasoning». En Habel, C., y Rickheit, G. (eds.), Mental Models in Discourse Processing and Reasoning. Nueva York: Elsevier Science, pp. 299-331.

TALL, N. (2006), Blank: the Power of Not Actually Thinking at All (A Mindless Parody). Nueva York: Harper Paperbacks.

TALMY, L. (1983), «How language structures space». En Pick, H., y Acredolo, L. (eds.), Spatial Orientation: Theory, Research, and Application. Nueva York: Plenum Press. (Reimpreso en una versión revisada y ampliada en Talmy, L. [2000], Toward a Cognitive Semantics, Vol. 1. Cambridge, MA: MIT Press, pp. 177-254.)

TAPLIN, J. E.; STAUDENMAYER, H., y TADDONIO, J. L. (1974), «Developmental changes in conditional reasoning: linguistic or logical?», Journal of Experimental Child Psychology, 17, 360-373.

TARSKI, A. (1944), «The semantic conception of truth», Philosophy and Phenomenological Research, 4, 341-375.

TARSKI, A. (1956), «The concept of truth in formalized languages». En Woodger, J. H. (ed.), Logic, Semantics, Metamathematics. Oxford: Oxford University Press, pp. 152-278. (Publicado originalmente en 1936.)

TARSKI, A. (1965), Introduction to Logic and to the Methodology of Deductive Sciences, 3ª ed. Nueva York: Oxford University Press.

TAYLOR, R. (1966), Action and Purpose. Englewood Cliffs, NJ: Prentice-Hall.

TEASDALE, J. D. (1988), «Cognitive vulnerability to persistent depression», Cognition and Emotion, 2, 247-274.

Tennant, N. (2005), «On the degeneracy of the full AGM-theory of theory-revision», Artículo de discusion, Universidad del Estado de Ohio.

TERTULIANO, Q. S. (1956), Treatise on the Incarnation (ed. y trad. Evans, E.), Londres: SPCK. (Una traducción de De Carne Christi.)

THAGARD, P. (1989), «Explanatory coherence», Behavioral and Brain Sciences, 12, 435-502.

THAGARD, P. (1992), Conceptual Revolutions. Princeton, NJ: Princeton University Press.

THAGARD, P. (1999), How Scientists Explain Disease. Princeton, NJ: Princeton University Press.

THAGARD, P. (2000), Coherence in Thought and Action. Cambridge, MA: MIT Press.

THAGARD, P., y HOLYOAK, K. J. (1985), «Discovering the wave theory of sound: inductive inference in the context of problem solving». En Proceedings of the Ninth International Joint Conference on Artificial Intelligence. Los Altos,

CA: Morgan Kaufmann, pp. 610-612.

THESIGER, W. (1959), Arabian Sands. Londres: Penguin. [Ed. cast.: Arenas de Arabia, trad. G. Rodríguez Clemente, Península, 2003.]

THODARSON, D. S., y SHAFRAN, R. (2002), «Importance of thoughts». En Frost, R. O., y Steketee, G. (eds.), Cognitive Approaches to Obsessions and Compulsions: Theory, Assessment, and Treatment. Boston, MA: Pergamon, pp. 15-28.

THOMPSON, V. A. (1994), «Interpretational factors in conditional reasoning», Memory & Cognition, 22, 742-758.

TOBIN, J. (2003), To Conquer the Air: the Wright Brothers and the Great Race for Flight. Nueva York: Free Press.

TODOROV, A.; MANDISODZA, A. N.; GOREN, A., y HALL, C. C. (2005), «Inferences of competence from faces predict election outcomes», Science, 308, 1623-1626.

TOGNOTTI, E. (2000), Il mostro asiatico. Storia del colera in Italia. Bari: Laterza.

TOLMAN, E. C. (1948), «Cognitive maps in rats and men», Psychological Review, 55, 189-208.

TOULMIN, S. E. (1958), The Uses of Argument. Cambridge: Cambridge University Press.

TRABASSO, T.; RILEY, C. A., y WILSON, E. G. (1975), «The representation of linear order and spatial strategies in reasoning: a developmental study». En Falmagne, R. (ed.), Reasoning: Representation and Process. Hillsdale, NJ: Erlbaum, pp. 201-229.

TRANEL, D. (1997), «Emotional processing and the human amygdala», Trends in Cognitive Sciences, 1, 46-47.

TRYON, W. W. (2005), «Possible mechanisms for why desensitisation and exposure therapy work», Clinical Psychology Review, 25, 67-95.

TULKIN, S. R., y KONNER, M. J. (1973), «Alternative conceptions of intellectual functioning», Human Development, 16, 33-52.

TURNBULL, W., y SLUGOSKI, B. R. (1988), «Conversational and linguistic processes in causal attribution». En Hilton, D. (ed.), Contemporary Science and Natural Explanation: Common-sense Conceptions of Causality. Brighton, Sussex: Harvester Press, pp. 66-93.

TVERSKY, A. (1977), «Features of similarity», Psychological Review, 84, 327-352.

TVERSKY, A., y KAHNEMAN, D. (1973), «Availability: a heuristic for judging frequency and probability», Cognitive Psychology, 5, 207-232.

TVERSKY, A., y KAHNEMAN, D. (1974), «Judgment under uncertainty: heuristics and biases», Science, 185, 1124-1131.

TVERSKY, A., y KAHNEMAN, D. (1980), «Causal schemas in judgments under uncertainty». En Fishbein, M. (ed.), Progress in Social Psychology. Hillsdale, NJ: Erlbaum, pp. 49-72.

TVERSKY, A., y KAHNEMAN, D. (1982a), «Judgments of and by representativeness». En Kahneman, D.; Slovic, P., y Tversky, A. (eds.), Judgement Under Uncertainty: Heuristics and Biases. Cambridge, Cambridge University Press, pp. 84-98.

TVERSKY, A., y KAHNEMAN, D. (1982b), «Causal schemas in judgements under uncertainty». En Kahneman, D.; Slovic, P., y Tversky, A. (eds.), Judgement Under Uncertainty: Heuristics and Biases. Cambridge, Cambridge University Press, pp. 117-128.

TVERSKY, A., y KAHNEMAN, D. (1983), «Extensional versus intuitive reasoning: the conjunction fallacy in probability judgment», Psychological Review, 90, 292-315.

TVERSKY, A., y KOEHLER, D. J. (1994), «Support theory: a nonextensional representation of subjective probability», Psychological Review, 101, 547-567.

TVERSKY, A., y SHAFIR, E. (1992), «The disjunction effect in choice under uncertainty», Psychological Science, 3, 305-309.

ULEMAN, J. S.; BLADER, S. L., y TODOROV, A. (2005), «Implicit impressions». En Hassin, R. R.; Uleman, J. S., y Bargh, J. A. (eds.). The New Unconscious. Nueva York: Oxford University Press, pp. 362-392.

ULLMAN, S. (1996), High-Level Vision: Object Recognition and Visual Cognition. Cambridge, A: MIT Press.

UNSWORTH, S. J., y MEDIN, D. L. (2005), «Cultural differences in belief bias associated with deductive reasoning?», Cognitive Science, 29, 525-529.

VADEBONCOEUR, I., y MARKOVITS, H. (1999), «The effect of instructions and information retrieval on accepting the premises in a conditional reasoning task», Thinking & Reasoning, 5, 97-113.

VALÉRY, P. (1958), «Poetry and abstract thought». En The Art of Poetry (trad. Folliot, D.). Princeton, NJ: Princeton University Press, pp. 52-81.

VALÉRY, P. (1970a), Analects (Trad. Gilbert, S.). Princeton, NJ: Princeton University Press.

VALÉRY, P. (1970b), «Moments». En Poems in the Rough (trad. Corke, H.). Londres: Routledge & Kegan Paul, pp. 75-83.

VALÉRY, P. (1972), «Literary reminiscences». En Leonardo Poe Mallarmé (trad. Cowley, M., y Lawler, J. R.). Londres: Routledge & Kegan Paul, pp. 317-324.

VAN DER HENST, J.-B.; YANG, Y., y JOHNSON-LAIRD, P. N. (2002), «Strategies in sentential reasoning», Cognitive Science, 26, 425-468.

VANDIERENDONCK, A., y DE VOOGHT, G. (1997), «Working memory constraints on linear reasoning with spatial and temporal contents», Quarterly Journal of Experimental Psychology, 50A, 803-820.

VANDIERENDONCK, A.; DIERCKX, V., y DE VOOGHT, G. (2004), «Mental model construction in linear reasoning: evidence for the construction of initial annotated models», Quarterly Journal of Experimental Psychology, 57A, 1369-1391.

VAN FRAASEN, B. (1973), «The logic of conditional obligation», Journal of Philosophical Logic, 1, 417-438.

VAPNIK, V. N. (1998), Statistical Learning Theory. Nueva York: Wiley.

VENDLER, H. H. (2004), Poets Thinking: Pope, Whitman, Dickinson, Yeats. Cambridge, MA: Harvard University Press.

VERHEIJ, B. (2003), «Artifitial argument asistants for defeasible arguments», Artifitiala Intelligence, 150, 291-324.

VERWEIJ, A. C.; SIJTSMA, K., y KOOPS, W. (1999), «An ordinal scale for transitive reasoning by means of a deductive strategy», International journal of Behavioral Development, 23, 241-264.

VERSCHUEREN, N.; SCHAEKEN, W., y D'YDEWALLE, G. (2005a), «Everyday conditional reasoning: a working memory-dependent tradeoff between counterexample and likelihood use», Memory & Cognition, 33, 107-119.

VERSCHUEREN, N.; SCHAEKEN, W., y D'YDEWALLE, G. (2005b), «A dual-process specification of causal conditional reasoning», Thinking & Reasoning, 11, 239-278.

VINTEN-JOHANSEN, P. V.; BRODY, H.; PANETH, N.; RACHMAN, S., y RIP, M. (2003), Cholera, Chloroform, and the Science of Medicine: a Life of John Snow. Oxford: Oxford University Press.

VON FRISCH, K. (1966), The Dancing Bees, 2ª ed. Londres: Methuen.

VON MISES, R. (1957), Probability, Statistics and Truth. Segunda edición inglesa revisada basada en la tercera edición alemana de 1951. Londres: Allen and Unwin.

VON NEUMANN, J. (1958), The Computer and the Brain. New Haven, CO: Yale University Press.

VOSNIADOU, S., y BREWER, W. F. (1992), «Mental models of the earth: a study of conceptual change in childhood», Cognitive Psychology, 24, 535-585.

VOSS, J. F.; FINCHER-KIEFER, R.; WILEY, J., y SILFIES, L. N. (1993), «On the processing of arguments», Argumentation, 7, 165-181.

WALSH, C., y JOHNSON-LAIRD, P. N. (2004), «Co-reference and reasoning», Memory & Cognition, 32, 96-106.

WALTON, D. N. (1989), Informal Logic: a Handbook for Critical Argumentation. Cambridge: Cambridge University Press.

WANG, H. (1963), «Dominoes and the AEA case of the decision problem». En Fox, J. (ed.), Mathematical Theory of Automata. Nueva York, NY: Instituto Politécnico de Brooklyn, pp. 23-56.

WASON, P. C. (1959), «The processing of positive and negative information», Quarterly Journal of Experimental Psychology, 11, 92-107.

WASON, P. C. (1960), «On the failure to eliminate hypotheses in a conceptual task», Quarterly Journal of Experimental Psychology, 12, 129-140.

WASON, P. C. (1966), «Reasoning». En Foss, B. M. (ed.), New Horizons in Psychology. Harmondsworth, Middx: Penguin, pp. 135-151.

WASON, P. C., y EVANS, J. St. B. T. (1975), «Dual processes in reasoning?», Cognition, 3, 141-154.

WASON, P. C., y JOHNSON-LAIRD, P. N. (1972), The Psychology of Deduction: Structure and Content. Cambridge, MA: Harvard University Press. Londres: Batsford.

WASON, P. C., y SHAPIRO, D. (1971), «Natural and contrived experience in a reasoning problem», Quarterly Journal of Experimental Psychology, 23, 63-71.

WATANABE, S.; SAKAMOTO, J., y WAKITA, M. (1995), «Pigeons' discrimination of paintings by Monet and Picasso», Journal of the Experimental Analysis of Behavior, 63, 165-174.

WATSON, J. S.; GERGELY, G.; CSANYI, V.; TOPAL, J.; GACSI, M., y SARKOZI, Z. (2001). «Distinguishing logic versus association in the solution of an invisible displacement task by children and dogs: Using negation of disjunction», Journal of Comparative Psychology, 115, 219-226.

WEISBERG, R. W. (1986), Creativity: Genius and Other Myths. Nueva York: Freeman.

WEISBERG, R. W. (1993), Creativity: Beyond the myth of genius. Nueva York: Freeman.

WEISBERG, R. W. (1996), «Prolegomena to theories of insight in problem solving: a taxonomy of problems». En Sternberg, R. J., y Davidson, J. E. (eds.), The Nature of Insight. Cambridge, MA: MIT Press, pp. 157-196.

WELCHMAN, G. (1982), The Hut Six Story: Breaking the Enigma Codes. Nueva York: McGraw-Hill.

WELLMAN, H. M.; CROSS, D., y WATSON, J. (2001), «Meta-analysis of theory of mind development: the truth about false-belief», Child Development, 72, 655-684.

WELLS, A. (2000), Emotions Disorders and Metacognition: Innovative Cognitive Therapy. Chichester, UK: Wiley.

WERTHEIMER, M. (1961), Productive Thinking. (Edición ampliada, ed. Michael Wertheimer.) Londres: Tavistock Press. (Publicado originalmente en 1945.)

WETHERICK, N. E., y GILHOOLY, K. J. (1990), «Syllogistic reasoning: effects of premise order». En Gilhooly, K.; Keane, M. T. G.; Logie, R., y Erdos, G. (eds.), Lines of Thought: Reflections on the Psychology of Thinking, Vol. 1. Londres: Wiley, pp. 99-108.

WHARTON, C. M., y GRAFMAN, J. (1998), «Deductive reasoning and the brain», Trends in Cognitive Sciences, 2, 54-59.

WHEWELL, W. (1967), The Philosophy of the Inductive Sciences, Founded Upon Their History, 2 vols. Londres: J. W. Parker. Facsimile de la 2ª ed. de 1847. Nueva York: Johnson Reprint.

WHITE, P. A. (1995), «Use of prior beliefs in the assignment of causal roles: causal powers versus regularity-based accounts», Memory and Cognition, 23, 243-254.

WHITEHEAD, A. N., y RUSSELL, B. A. W. (1927), Principia Mathematica, 2ª ed. Cambridge: Cambridge University Press.

WHITMAN, W. (1977), The Complete Writings of Walt Whitman. St. Glair Shores, MI: Scholarly-Press.

WHYTE, L. L. (1978), The Unconscious Before Freud. Nueva York: St. Martin's Press.

Wilson, T. D. (2002), Strangers to Ourselves: Discovering the Adaptive Unconscious. Cambridge, MA: Belknap, Harvard University Press.

WIMMER, H., y PERNER, J. (1983), «Beliefs about beliefs. Representation and containing function of wrong beliefs in young children's understanding of deception», Cognition, 13, 103-128.

WINKELMANN, J. H. (1980), «Semantic representation of kinship systems», Journal of Psycholinguistic Research, 4, 133-140.

WINSTON, P. H. (1975), «Learning structural descriptions from examples». En Winston, P. H. (ed.), The Psychology of Computer Vision. Nueva York: McGraw-Hill, pp. 157-209.

WISE, M. N. (1979), «The mutual embrace of electricity and magnetism», Science, 203, 1310-1318.

WITTGENSTEIN, L. (1922), Tractatus Logico-Philosophicus. Londres: Routledge & Kegan Paul. [Ed. cast.: Tractatus Logico-Philosophicus, trad. de L. M. Valdés Villanueva, Barcelona, Península, 2003.]

WITTGENSTEIN, L. (1953), Philosophical Investigations (trad. Anscombe, G. E. M.). Nueva York: Macmillan.

WOLPE, J. (1990), The Practice of Behavior Therapy. Nueva York: Pergamon.

WOOD, D. J. (1969), «Approach to the study of human reasoning», Nature, 223, 102-103.

WOOD, D. J., y SHOTTER, J. D. (1973), «A preliminary study of distinctive features in problem solving», Quarterly Journal of Experimental Psychology, 25, 504-510.

WOOD, D. J.; SHOTTER, J. D., y GODDEN, D. (1974), «An investigation ot

the relationships between problem solving strategies, representation and memory», Quarterly Journal of Experimental Psychology, 26, 252-257.

WOODS, M. (1997), Conditionals. Oxford: Clarendon Press.

WOODWORTH, R. S. (1938), Experimental Psychology. Nueva York: Holt.

WOODWORTH, R. S., y SELLS, S. B. (1935), «An atmosphere effect in formal reasoning», Journal of Experimental Psychology, 18, 451-460.

WOTIZ, J. H., y RUDOFSKY, S. (1984), «Kekule's dreams: fact or fiction?», Chemistry in Britain, 20, 720-723.

WRIGHT, O. (1988), How We Invented the Airplane: an Illustrated History. Nueva York: Dover. (Publicado originalmente en 1953.)

YANG, Y., y JOHNSON-LAIRD, P. N. (2000a), «Illusory inferences with quantified assertions: how to make the impossible seem possible, and vice versa», Memory & Cognition, 28, 452-465.

YANG, Y., y JOHNSON-LAIRD, P. N. (2000b), «How to eliminate illusions in quantified reasoning», Memory & Cognition, 28, 1050-1059.

ZAJONC, R. B. (1980), «Feeling and thinking: preferences need no inferences», American Psychologist, 35, 151-175.

ZHANG, J.; JOHNSON, T. R., y WANG, H. (1997), «The relation between order effects and frequency learning in tactical decision making», Thinking & Reasoning, 4, 123-145.

ZWAAN, R. A., y RADVANSKY, G. A. (1998), «Situation models in language comprehension memory», Psychological Bulletin, 123, 162-185.